

DETECTION OF PECAN WEEVIL LARVAE  
IN PECAN NUTMEAT USING  
MULTISPECTRAL IMAGING SYSTEM

By

CHAITAL P. SHAH

Bachelor of Engineering in Mechatronics,

Sardar Patel University,

G.H.Patel College of Engineering and Technology,

Vallabh Vidyanagar, India

1998

Submitted to the faculty of the  
Graduate College of the  
Oklahoma State University  
in partial fulfillment of  
the requirement for  
the Degree of  
MASTER OF SCIENCE  
July, 2006

DETECTION OF PECAN WEEVIL LARVAE  
IN PECAN NUTMEAT USING  
MULTISPECTRAL IMAGING SYSTEM

Thesis Approved:

Dr. Paul Weckler

---

Thesis Adviser

Dr. Marvin Stone

---

Dr. Neils O Maness

---

Dr. Timothy Bowser

---

Dr. A. Gordon Emslie

---

Dean of the Graduate College

## **Acknowledgements**

This thesis is a result of one and half years of work whereby I have been accompanied and supported by many people. It is a pleasant aspect that I have now the opportunity to convey my gratitude for all of them.

First and foremost I am deeply grateful to my thesis advisor and mentor Dr. Paul Weckler, for his detailed and constructive comments and for his important support throughout this work. I would also like to thank him for providing me with the resources that were required for this research. He taught me how to write academic papers. More importantly, he taught me how to work hard and without his encouragement and constant guidance, I could not have finished this dissertation. He was always there to meet and talk about my ideas, to proofread and mark up my papers and chapters, and to ask me good questions to help me think through my problems.

My appreciation goes to my other committee members as well. I thank Dr. Marvin Stone for the technical discussions on the spectral response measurements and for the help with Machine vision Instrumentation. Our numerous scientific discussions and his many constructive comments have greatly improved this work. It was my great pleasure to interact with Dr Niels O. Maness from the department of Horticulture and Landscape Architecture. He taught me how to ask questions and express my ideas. He showed me different ways to approach a research problem and the need to be persistent to accomplish a goal. He provided me with the resources that were

necessary for the research. Thanks to all of you not just for being on my committee but for all of your guidance and assistance.

I am also greatly indebted to Dr. Guoliang Fan, Department of Electrical and Computer Science Engineering, for getting me interested in image processing and computer vision. He taught me fundamentals of signal and image processing and showed constant interest in my research work.

I would also like to greatly acknowledge the support of some very special individuals. Madhuri Cumbamanagalam helped me immensely by giving me encouragement and friendship. She had confidence in me when I doubted myself, and brought out the good ideas in me. She helped me with proof reading the thesis chapters and making grammatical corrections wherever necessary.

I thank my parents for educating me with principles of life and science. They gave me unconditional support and encouragement to pursue my interests and providing me with countless opportunities for which I am eternally grateful.

# Table of Contents

Chapter	Page
<b>1.0 Introduction</b>	
1.1 Pecan [ <i>Carya illenoinensis</i> (Wangenh.) K. Koch].....	1
1.2 Pecan weevil [ <i>Curculio caryae</i> ].....	2
1.3 Objectives .....	5
<b>2.0 Literature Review</b>	
2.1 Introduction.....	6
2.2 Spectroscopy .....	7
2.2.1 Visible and near infrared spectroscopy .....	7
2.2.2 Diffuse reflectance spectroscopy .....	8
2.2.3 Parameters effecting spectral measurements.....	9
2.3 Noise and dynamics in spectral measurements .....	11
2.4 Spectral data analysis .....	13
2.4.1 Multivariate statistical analysis .....	13
2.4.2 Derivative analysis .....	14
2.5 Imaging techniques .....	15

2.5.1 Multispectral imaging .....	16
2.5.1.1 Instrumentation for multispectral imaging system.....	16
2.6 Image processing .....	18
2.6.1 Threshold techniques .....	18
2.6.2 Edge-based methods .....	18
2.6.3 Region based methods .....	19
2.6.4 Mixed methods that combine edge and region based methods....	19

### **3.0 Materials and methods**

3.1 Diffuse reflectance measurement .....	21
3.2 Analyzing spectral measurement data .....	24
3.2.1 Principal component analysis and derivative analysis .....	25
3.3 Selection of narrow bandpass interference filter and NIR enhanced camera .....	27
3.4 Acquiring and processing image .....	32
3.4.1 Components of an image processing system .....	32
3.5 Image Segmentation .....	33
3.5.1 Method 1 .....	34
3.5.1.1 Canny edge detector .....	34
3.5.2 Method 2 .....	36
3.5.2.1 Active contour based snake energy minimization algorithm .....	36
3.5.2.2 Region growing .....	39
3.5.2.3 Problems with traditional seeded region growing And solution .....	40
3.5.2.4 Region merging .....	42

3.5.2.5 The problem faced with this algorithm .....	42
<b>4.0 Results and discussions</b>	
4.1 Diffuse reflectance measurements.....	44
4.2 Data analysis.....	46
4.3 Image acquisition and Canny edge detection.....	50
4.4 Segmentation results for training images using Method 1.....	52
4.5 Segmentation results: Type 2 error.....	58
4.6 Segmentation results for testing images using Method 1.....	63
4.7 Segmentation result using Method 2.....	68
<b>5.0 Conclusions and future work.....</b>	<b>74</b>
<b>6.0 References .....</b>	<b>77</b>

## List of Tables

Table	Page
3.1 Specifications of Labspec Pro spectrometer .....	24
3.2 Specifications of narrow band-pass interference filters .....	29
3.3 Specifications of AVT Dolphin® camera .....	31
4.1 Segmentation results for 50 training images .....	56
4.2 Segmentation results determining Type 2 error.....	61
4.3. Segmentation results for 50 testing images .....	66
5.1 Classification results .....	75



## List of Figures

Figure	Page
1.1 Hole drilled by the larvae as it exits the nut .....	3
1.2 Damaged pecan nuts .....	3
1.3 Damaged interior of the nut .....	4
3.1 A typical configuration for a reflection experiment .....	23
3.2 Narrow band pass interference filter from Edmund Optics .....	30
3.3 Spectral response of a typical silicon CCD .....	31
3.4 Spectral response of the AVT Dolphin 145-B camera .....	32
3.5 Components of an image processing system .....	33
3.6 Structuring element .....	37
3.7 Complete Algorithm flowchart .....	38
3.8a Rectangular boundary .....	42
3.8b Random boundary .....	42
4.1 Diffuse reflectance measurements of pecan weevil larvae and nutmeat - 1 .....	44
4.2 Diffuse reflectance measurements of pecan weevil larvae and nutmeat - 2 .....	45
4.3 Diffuse reflectance measurements of pecan weevil larvae and nutmeat - 3 .....	45
4.4 Average diffuse reflectance of weevil larvae, weevil larvae preserved in 95% ethanol and pecan nutmeat.....	46
4.5 Principal component analysis .....	47
4.6 First derivative of nutmeat .....	48
4.7 First derivate of weevil larvae .....	48
4.8 First derivative of preserved weevil larvae .....	49
4.9 Derivative graphs superimposed .....	49

4.10 Example image acquired at 980 nm .....	51
4.11 Gaussian filtered image .....	51
4.12 Result of Canny edge detection .....	52
4.13 Segmentation results of 50 training images using method 1 .....	53
4.14 Image threshold .....	56
4.15 Segmentation results: Type 2 error .....	59
4.16 Segmentation results for testing images using method 1 .....	63
4.17 Dilation residue edge detector .....	68
4.18 Deployment of a seed point .....	69
4.19 Region growing, iteration 10 - 60 .....	70
4.20 Energy minimization .....	73

## Nomenclature

$\oplus$	--	dilation
$\bullet$	--	closing
$\ominus$	--	erosion
$\circ$	--	opening
$\theta$	--	noise (stray light and dark current)
A	--	absorbance
ADC	--	analog to digital converter
ADU	--	analog digital unit
AI	--	absorbance index
AOTF	--	acoustic optic tunable filter
$A_r$	--	apparent absorbance (in reflectance mode)
$A_\lambda$	--	absorbance at wavelength $\lambda$
B/W	--	black and white
BLPF	--	Butterworth lowpass filter
BTH	--	black top hat transformation
CMYK	--	cyan, magenta, yellow and black
CVF	--	circular variable filter
CWL	--	central wavelength
D	--	effective dynamical range

$D_\lambda$	--	dark intensity of each wavelength
$E_{\text{cost}}$	--	cost energy function
eig1	--	eigen vector 1
EM	--	expectation maximization
FIR	--	finite impulse response
FPN	--	fixed pattern noise
FWHM	--	full-width half- maximum
$I_a$	--	absorbed radiation
ILPF	--	ideal low pass filter
$I_o$	--	incident radiation
$I_r$	--	reflected radiation
$I_t$	--	transmitted radiation
LCTF	--	liquid crystal tunable filter
LoG	--	Laplacian of Gaussian
Mb/s	--	mega bits per second
nm	--	nanometer
PC1	--	first principle component
PC2	--	second principle component
PCA	--	principle components analysis
PDF	--	probability density function
PET	--	positron emission tomography
R	--	measured reflectance
$R_{\text{ADC}}$	--	resolution of A/D converter

S:N	--	signal to noise ratio
SRG	--	seeded region growing
$S_{\lambda}$	--	sample intensity of each wavelength
TE	--	thermo electric
USDA	--	United States Department of Agriculture
WTH	--	white top hat transformation
$\mu\text{m}$	--	micrometer

## **1. Introduction**

### **1.1 Pecan [*Carya illenoinensis* (Wangenh.) K. Koch]**

Pecan, *Carya illenoinensis* (Wangenh.) K. Koch is a tree nut crop that largely grows in deciduous forests. United States is the world's largest producer of pecan followed by Mexico. Pecan and black walnut are the only native nut crops grown in United States. Pecan is the most important commercial crop with high consumer value. The commercial significance of black walnut is limited. (Economic Research Service, USDA, May 2003). Southern and southwestern states such as Texas, Georgia and New Mexico are the leading producer of pecan nuts. The other major producers within United States include Alabama, Louisiana and Oklahoma. The other principal producing countries are Australia, Brazil, Israel, Peru and South Africa.

Pecan production can be classified into two categories: *native and improved variety*. The improved variety is superior to the native seedlings and requires intensive cultivation practices. In year 2002, the total acreage of pecan was estimated at over 1.4 million acres. Out of this 60 % were native seedlings while 40% were improved variety (USDA, 2002). Oklahoma's pecan industry depends largely on the native seedlings with ninety percent of the total production is native seedlings. Its annual average production is about 8 million kilograms of the total 104 million kilograms produced in Unites States. As pecan exhibits alternate bearing, production in an individual year can range from 1 million up to 9 million kilograms with an occasional year exceeding 14 million

kilograms. Oklahoma has the third highest number of acres planted to pecans. Its production, however, ranks about fifth place among other states because its crop is heavily dependent on native seedlings (McCraw, 2004).

Pecans are a favorite choice among consumers because of their outstanding taste. They add flavor, crunch and texture to various foods or snacks. They are most widely used in bakery, confectionary, chocolate and ice cream industries (Saaristo, 2005). The discarded nuts can be used to extract oils that are often used in cooking and cosmetics. Pecan milk is made from the seed that is habitually used to season corn cakes and thicken soups (Facciola, 1990). Pecans also have medicinal value. They are used for the treatment of ringworms (Moerman, 1998).

## **1.2 Pecan weevil [*Curculio caryae*]**

There are numerous categories of insects that affect pecan nut crop. Aphids, mites and pecan phylloxera are the most common type of insects that attack the pecan leaf. They are generally known as leaf feeders. Shot-hole borer, twig girdler and flat-headed apple tree borer are referred to as trunk and twig feeders. Pecan weevil, shuck worm, pecan nut case bearer, stinkbugs and spittlebugs are nut feeders and directly damage the pecan nuts (Bessin, 2005).

Pecan weevils are the most serious late season pests as they affect the nut. It attacks both the native and improved varieties. Adult pecan weevils are 1 cm long brownish beetles with an extremely long beak. Pecan weevil grubs have no legs; they are creamy white with reddish-brown heads. They grow up to 2 cm long (Douce, 2000). The life cycle of pecan weevil is between two and three years. Adult pecan weevils typically emerge from the soil in the month of August. Emergence of pecan weevil is usually after

two to three days of heavy rains. Factors such as rainfall, soil moisture and type of soil affect the ability of the weevils to emerge from the soil (Chaney, 2005). For example, a brief period of drought can holdup the emergence of adult pecan weevil. After the nut kernel hardens, the pecan nut becomes suitable for egg laying. The female weevil drills a hole through the shell using her long snout. It then lays eggs in the developing kernel. Larvae feed on the developing nut for about 35 days damaging the nut. After which they return to the soil. The figure 1.1 shows the hole drilled by the larvae as it exits the nut. They remain in larvae stage for about one or two years after which they pupate.



**Figure 1.1 Hole drilled by the larvae as it exits the nut**

Pecan weevils can induce two types of damage depending on the stage of nut development. The adult weevils feed on the nut during the water stage prior to nut hardening. This causes nuts to wither, turn black and ultimately drop. These damaged nuts are shown in the figure 1.2.



**Figure 1.2 Damaged pecan nuts**



This damage is not easy to distinguish and may be confused with damage from sucking bugs or the shuck worm. Similar damage can also be seen due to a mechanical injury. The pecan weevil larvae are only seen in a hardened shell. If the larvae are seen prior to shell hardening, then the damage can be potentially from some other insect, probably shuck worm (Douce, 2000). Nuts damaged by the pecan weevil have pin-sized holes penetrating deep inside the nut.

The other damage is more serious. It is caused by weevil larvae feeding on developing kernels. They feed on the kernels for four to five weeks, tearing down the interior of the nut as shown in figure 1.3. Damaged nuts can often be recognized by circular emergence holes usually 3 mm in diameter, through which the grubs exit ( Ellis, 2005).



**Figure 1.3 Damaged interior of the nut**

### **1.3 Objectives**

The overall objective of this study was to design a multispectral imaging system for detecting pecan weevil larvae in pecan nuts during post-harvest processing.

Hyperspectral data analysis was used to determine the wavelengths for selection of specific band pass filters. The four specific steps to achieve the objective were as follows:

1. Obtain hyperspectral data of pecan weevil larvae and pecan nutmeat using reflectance and/or transmission spectroscopy (450nm- 1015nm).
2. Analyze the hyperspectral data to detect the pecan weevil larvae in pecan nuts.
  - a) Identify two to four wavelengths that can potentially lead to selection of band pass filters.
3. Implement a multispectral imaging system by using the identified wavebands.
4. Develop and test a multispectral imaging algorithm to classify the pecan weevil larvae in pecan nutmeat.

## **2. Literature Review**

### **2.1 Introduction**

Optical spectroscopy offers fast and nondestructive means to assess agricultural commodities for their quality and safety for human consumption. Optical spectroscopy is remarkably adaptable and robust analytical technique that is gaining popularity due to its accuracy, efficiency and ease of use (Roberts, Workman and Reeves, 2004). Spectral sensing techniques are commonly used for inspection of food commodities. This non-invasive technique commonly uses spectral regions ranging through the visible (VIS) to the near infrared (NIR). A number of papers, journal articles and conference proceedings have demonstrated the application of this technique to the detection of defects in apples as well as its physical properties such as size and color (Miller et al., 1998; Upchurch et al., 1990; Upchurch et al., 1994; Tao, 1997; Throop et al., 1995), spectral image analysis for measuring ripeness of tomatoes (Podler et al., 2002), single kernel maize analysis by near infrared hyperspectral imaging (Cogdil et al., 2004), automated sorting of almonds with embedded shell by laser transmittance imaging (Pearson et al., 1995), machine vision system for automated detection of stained pistachio nuts (Pearson, 1995) and detection of internal insect infestation in whole wheat kernels (Brabec and Schwartz, 2003).

Currently, multispectral imaging is widely used in produce grading, meat inspection and defect detection. Defect detection is accomplished by imaging at specific wavelengths where the reflectivity of the good and the damaged tissue are different. Quality control is achieved by classification of a particular attribute, like color, which can determine the grade of the product. The changes in spectral reflectivity can be used to detect the presence of toxins in fish and poultry. In addition to the above applications, multispectral imaging is also used in foreign body detection, in which the difference in reflectivity between the product and the foreign body is used for detection (Park, 1996).

## **2.2 Spectroscopy**

Spectroscopy is the interaction of electromagnetic radiation with matter. It is a method used to identify or quantify molecular structure or to detect physical properties of matter by studying and examining interactions of electromagnetic radiation with molecules of the matter.

### **2.2.1 Visible and near infrared spectroscopy**

The wavelength range of radiation covered in optical spectroscopy can be divided into Ultraviolet (UV), 200 to 400 nm; Visible, 400 to 700nm, Near Infrared (NIR), 700 to 1500 nm, Short-Wave Infrared (SWIR), 1.5 $\mu$ m to 3 $\mu$ m and Mid-Wave Infrared (MWIR), 3 $\mu$ m to 5.5 $\mu$ m. The Long Wave Infrared (LWIR) waveband starts at about 8 $\mu$ m and extends to about 14 $\mu$ m.

The recent trend of application of NIR spectroscopy in the field of food and agriculture is largely due to the work done by Norris. et al. (1968). Norris recognized that food and agriculture material exhibit C-H, N-H and O-H absorption bands identifiable in the NIR region of the electromagnetic spectrum. For the visible region of the spectrum

the term absorbance and reflectance is easy to comprehend. The red color looks red as it reflects red color and absorbs all the other colors. The object looks white as it reflects all wavelengths and looks black as it absorbs all wavelength of visible light. The similar behavior is exhibited by NIR absorption. C-H, O-H and N-H absorption bands have specific vibrational frequency that results in distinctive NIR response for different biological materials (Shenk and Workman 2001). As many natural biological substances contains thousands of different types of compounds containing C-H, O-H and N-H molecular bonds, the absorption spectra in the NIR region is very complex to interpret. This complexity is further augmented by the fact that this spectral information is crammed in a very small region of the electromagnetic spectrum (Shenk and Workman 2001).

Apart from the chemical composition of biological materials, the physical texture and surface characteristics of the material also play a significant role in shaping the spectral response of the sample. Every biological substance has a unique spectral response. If two samples have similar spectral response, it can be concluded that they have similar physical and chemical compositions. If they have different spectral responses then the two samples are considered to be different physically or chemically. Therefore, the NIR absorbance or reflectance spectrum obtained from a substance is a result of the combination of all the chemical and physical information of the biological material (Roberts, Workman and Reeves, 2004).

### **2.2.2 Diffuse reflectance spectroscopy**

An electromagnetic radiation interacts with a sample through absorbance, reflectance and transmission. Some of the radiations are returned either by the boundary

of the sample or the interior of the sample. They are referred to as *surface reflectance* and *volume reflectance* respectively. Passage of electromagnetic radiation through the sample is known as transmission. Radiation that is not reflected or transmitted is absorbed by the sample. Absorbance is the ability of a substance to transform radiant energy into a different form, usually resulting in rise of temperature.

Mathematically, the interaction of electromagnetic radiation with a sample can be represented as:

$$I_o = I_r + I_a + I_t \quad (2.1)$$

Where  $I_o$  is incident radiation,  $I_a$  is absorbed radiation,  $I_r$  is reflected radiation and  $I_t$  is transmitted radiation.

Absorbance, reflectance and transmission depend on wavelength, polarization and geometric distribution of the incident radiation. Reflectance can further be classified into specular reflectance and diffuse reflectance. If the angular distribution of the reflected energy is independent of the incident angle of radiation, it is known as diffuse reflectance. Therefore, only that part of radiation that is scattered inside the sample and reflected or the radiations that are reflected from a rough surface are measured to be diffuse reflectance radiations. The sample may have both diffuse and specular reflectance (Roberts, Workman and Reeves, 2004).

### **2.2.3 Parameters effecting spectral measurements**

There are various factors that have significant effect on the strength and noise of the measured spectral signal. Integration time, dynamic average of samples, average of pixels, dark current and offset are some of the major factors.

The signal that is displayed is proportional to the number of photons that strike the detector. A longer *integration time* can be used to increase the signal intensity level to a sufficient count. The integration time of the spectrometer is analogous to the shutter speed of a camera. The higher the value specified for the integration time, the longer the detector collects the incoming photons. If the intensity of the acquired spectra is too low, the integration time needs to be increased and if the intensity level of the spectra is close to the saturation level then the integration time needs to be reduced. However, the influence of noise increases with the integration time.

One of the important parameters that needs to be adjusted is the *number of sample scans to be dynamically averaged*. Each spectrum is collected and the running average is calculated. The signal to noise ratio (S: N) will improve by the square root of number of scans averaged. The advantage of higher S: N ratio has to be traded with the amount of time required to scan. In most of the real-time applications, the number of scans that has to be averaged must be minimized. Identifying the appropriate number of scans is an important task.

Another important parameter that is the *number of pixels (adjacent detector elements) that have to be averaged*. The greater this value, the smoother the data and the higher the S: N ratio. However, if the value entered is too high it results in the loss of spectral resolution. As a consequence the averaging of pixels should be done after the identification of wavebands. The loss of spectral resolution is higher in non-weighted averaging techniques than in weighted averaging techniques.

One of the weighted averaging techniques that is most often used is called Savitzky- Golay averaging (Press et al., 2005). It is basically a FIR low pass filter and

can be used with frequency or spectroscopic data. The coefficients of this filter are determined by performing an unweighted linear least square fit using a polynomial of a given degree. It is used to determine the smoothed value for each data point by performing polynomial regression and hence is also known as a digital smoothing polynomial filter. A higher level of smoothing without attenuation of data can be achieved when a higher degree of polynomial is taken. A different polynomial is applied to the data surrounding each data point. The values of the smoothed points are calculated by replacing each data point with that of its fitted polynomial. A linear filter is usually used for calculating the coefficients of the polynomial (Press et al., 2005).

The *dark current* is the signal that is present without any light input. This is typically very small. The dark current comes from the detector itself i.e. it is inherent in the detector. As a consequence it contributes to the signal level of the measured pixel (Bolton, 2002). Cooling the detector array is the most effective way to reduce the dark current signal.

### **2.3 Noise and dynamics in spectral measurements**

Different pixel elements of the CCD array respond differently to the same light intensity. This results in noise that varies about some nominal value. In order to minimize the instrument noise, the measurements should be made only after the instrument is fully warmed up. It is necessary to ensure that the temperatures of the instrument and the detector do not fluctuate.

There are two main categories of sources, which generate noise in a line array, Fixed pattern noise (FPN) and Random noise.



**Fixed pattern noise (FPN)** is the variation in output pixel values, under uniform illumination. It results from different sensitivity to light (this includes differences in signal collection, amplification and digitalization) for different detector elements. When the read out conditions are set and the integration time is not changed the FPN is unvarying. FPN can be removed completely by subtracting the noise from the measured signal.

**Random noise** is the response variation for the same pixel. It is primarily due to the fluctuations in sensitivity of the same pixel under varying light intensities. The idea is to subtract the fixed pattern noise from the total noise. The secondary resultant noise is the random noise that determines the lower limit of the system dynamic range (Holtzman, 2005).

The dynamic range is the ratio of the highest signal (overexposure level) to the lowest signal (darkest) that a detector array can record. The highest signal corresponds to the brightest signal level just below saturation level and the darkest signal refers to the level below which no detail is perceived. The highest signal measurable is limited by the saturation level of the detector pixel elements and the lowest signal by the random noise of the spectroscopic system.

The one other source of noise or error in spectroscopic measurement is **stray light**. It refers to the radiation of the wrong wavelength that strikes the pixel. Stray light is responsible for most errors in low light measurements. Sources of stray light can be ambient light, imperfections of the grating, dust, scattering of light from imperfect optical components or reflections of non-optical components such as the spectrometer housing (Avantes, 2004).

Stray light influences the precision of the system in a critical way. It causes non-linearity of the signal at lower power levels and thus limits the measuring range of the system. One of the ways to overcome stray light noise is to limit the light to the region of the spectra. A more robust approach is to determine the stray light correction constant. This constant compensates for the total amount of stray light in the spectrometer. Both the approaches should be used in combination to overcome the influence of the stray light.

## **2.4 Spectral data analysis**

The goal of spectral data analysis is to identify the wavebands that capture maximum variance between pecan nutmeat and pecan weevil larvae. These wavebands are later used in designing a multispectral imaging system.

### **2.4.1 Multivariate statistical analysis**

Multivariate statistical analysis is a powerful spectrum data analysis tool. It provides a quantitative and robust method to simultaneously analyze high dimensional spectral data (Anderson, 2005). There are numerous standard methods of multivariate statistical analysis. Some of the widely used methods for analyzing spectral data are multiple (linear) regression, factor analysis, principle components analysis (PCA), discriminant analysis and cluster analysis.

Principal component and factor analysis are primarily used to reduce the number of variables in the data set or to classify variables by detecting the structural relationships between variables (StatSoft Inc., 2003). These are used for identifying patterns in data and expressing the data so as to highlight their similarities and differences. Principal component analysis has a wide range of applications. It is used in image and data

compression, noise filtering, color representation or visual interpretation with multi-band data, classification and recognition problems, computer science and biomedical signal processing (Japan Association of Remote Sensing, 1996).

Principle component analysis is a procedure of formulating new variables by using the information from the old variables. These new variables capture the maximum variance in the data set. The principle component is the linear combination of variables such that it extracts the maximum variance from the variables. This forms the first principle component (*PC1*). The second principle component (*PC2*) is the second linear combination of the variables that extracts the maximum proportion from the remaining variance. The *PC1* and *PC2* form the major and minor axis for the new set of data. These components are orthogonal as they are uncorrelated. The *PC1* and *PC2* provide two-dimensional visualization of the multidimensional data set (M&M Backgrounder Winter, 2005).

#### **2.4.2 Derivative analysis**

Derivatives of the spectral data are used to remove the constant background signals and improve the visual resolution of the spectroscopic data. The weak peaks that are not easily visible in the original spectral data can easily be identified by using the derivatives of the spectral data.

The first derivative of the spectrum is the slope of the spectrum. As the slope of a constant is zero, the first derivative effectively removes the constant background signals that offset the spectral data. The second derivative of the spectrum is the measure of the curvature of the spectrum. As the curvature of the linear slope is zero, the second derivative effectively removes the linear tilt of the spectral data. However, the second

derivative is sometimes hard to interpret as it changes the maxima into minima and vice versa (Siesler, Ozaki, Kawata and Heise, 2002).

There are several methods of calculating the derivatives of the spectral data. The commonly used methods are Fourier transform, Savitzky- Golay and finite difference method (Williams and Norris, 2001). The derivative graphs are sometimes hard to interpret; however the derivative analysis can sometimes give excellent results. The point at which the derivative changes sign is usually the point of interest as it indicates maximum variability between the spectral data sets.

## **2.5 Imaging techniques**

Imaging techniques can be classified into 4 methodologies.

1. ***Conventional imaging***: A single broadband filter with either B/W or a color imaging technology can be used.
2. ***Spectral imaging***: A single waveband is used and is typically implemented by using single camera and a narrow band pass interference filter.
3. ***Multispectral imaging***: It uses 2 – 20 wavebands and is implemented by using a single camera and a filter wheel or multiple cameras with an individual narrow band pass interference filter.
4. ***Hyperspectral imaging***: It uses from one hundred to a few hundred spectra and is implemented by using hyperspectral imagers or imaging spectrometers. A single camera coupled with liquid crystal tunable filter (LCTF), acoustic- optic tunable filter (AOTF) or a circular variable filter (CVF) is normally used (Park and Lawrence, 2005). In hyperspectral imaging, full spectral profile is coupled with each image pixel. As a consequence the resultant data quantities are excessively large. Substantial

amount of time is required to acquire hyperspectral images and to analyze huge amounts of data. This makes the online implementation of the hyperspectral imaging system almost impossible (Mehl, Chao, Kim and Chen).

### **2.5.1 Multispectral imaging**

Multispectral imaging is not a new concept. It has been extensively used over the years. The RGB images are the most common multispectral images. They contain three spectral color planes: *red, green and blue*. All the three spectral planes lie in visible portion of the electromagnetic spectrum. Another example of multispectral images are the CMYK (*Cyan, magenta, yellow and black*) images of color printing. When working with real time processing and inspection systems, speed requirements impose rigorous constraints on hardware for the imaging and optical systems. As a result, the hyperspectral system cannot be directly implemented on an inspection or a processing line. Under these circumstances, the multispectral imaging technology becomes a favorable approach. Multispectral image analysis is a faster technique as it is based on discrete spectral data analysis at a few selected wavelengths as opposed to the continuous spectral data analysis performed by hyperspectral imaging systems.

The analysis of spectroscopic or hyperspectral data leads to the selection of band pass interference filters. However transferring these results to multispectral equipment is limited to the spectral range of each channel available in the different cameras.

#### **2.5.1.1 Instrumentation for multispectral imaging system**

There are various factors that influence the selection of a digital camera. Some of the important factors are spatial resolution, spectral responsivity, quantum efficiency,

frame rate, shutter speed, availability of gain control, internal memory, data transfer rate, type of digital interface and the availability of external trigger.

The spectral sensitivity and quantum efficiency are the two important factors that must be considered while selecting a digital camera for multispectral applications.

Spectral sensitivity is the relative response of the light sensitive element (CCD) to various wavelengths in the electromagnetic spectrum while quantum efficiency is the effectiveness of the CCD to convert the incoming photon to electrical signal at a particular wavelength. Both spectral sensitivity and quantum efficiency are dependent on the material of the CCD elements. The choice of narrow band pass interference filter (waveband) that is used for acquiring images usually governs the selection of CCD material. For example, if a filter with waveband in the IR region of an electromagnetic spectrum is used then Indium Gallium Arsenide (InGaAs) CCD detectors are needed. InGaAs detectors are extremely sensitive to light in the wavelengths from 900 nm to 1700 nm. On the other hand, the silicon CCD is sensitive in the visible region of 400 nm to 700 nm and low NIR wavelength region of 700 nm to 1100 nm of the electromagnetic spectrum.

Once the definite wavebands are identified through spectral data analysis, the *band pass interference filters* are selected. The band pass interference filters transmit a particular band of frequencies while rejecting all the other upper and lower frequencies of an electromagnetic spectrum. The band pass interference filters can be classified as *narrow band pass* and *broad band pass interference filters*. The broad band pass interference filters pass a broadband of light. For instance, they transmit light from 375nm to 425 nm while blocking all the other wavebands. On the other hand, the narrow

band pass interference filter transmits a narrow band of light as little as 1 nm (Turner designs, 2005). The quality of a narrow band pass interference filter depends on various factors such as peak transmittance at central wavelength, FWHM (full-width half-maximum) and upper and lower cut-off frequencies.

## **2.6 Image processing**

Image segmentation and object recognition are the most important components of an image processing application. The principle behind the segmentation is to eliminate the useless or less useful information from an image and highlighting the useful features of an image. In segmentation, the pixels that have similar features are treated as a common class. There are two popular approaches to image segmentation. One of the most widely used methods is K-means segmentation and the second is Expectation Maximization (EM) algorithm.

Image segmentation algorithms falls under 4 categories.

### **2.6.1 Threshold techniques**

The most primitive method of image segmentation was based on thresholding techniques. The basic idea behind this technique was that the pixels whose value (gray intensity, color or other) is between two values of thresholds belong to one region. The values of the thresholds were derived from the histogram. This method did not take into consideration, any spatial information of the image. Therefore it failed to cope up with noise as well as with blurred edges of the image.

### **2.6.2 Edge-based methods**

In the early 1990s, people came up with several algorithms based on edge based image segmentation. These algorithms are considered to be powerful as they use the

method that closely corresponds to the human visual system. In this method, we find the places of rapid transition from one region to the other region of different brightness or color value. Gradient operators are convolved with the image. Various gradient operators such as Sobel, Herwitt and Roberts were developed during this period. Canny edge detector became very popular and is widely used even today. The basic drawback of this method is that the edges have to be linked to form closed boundaries of the regions. To go from image edges to the image boundaries is very difficult task, which has been studied a lot. The Canny edge detector is implemented in this research. The detailed description of the Canny edge detector is presented in chapter 3.

### **2.6.3 Region based methods**

In the late 1990s, several region based image segmentation algorithms were introduced. In this method the pixels of the same or similar brightness or color are grouped together to the regions according to the given criteria of homogeneity. The seeded region growing (SRG) is the most popular and widely used method. In this method, seed points are set up and the neighboring pixels that have properties similar to the seed point are appended to it. The criterion is based on some threshold value. The choice of threshold is problematic, as selecting a right threshold is usually very difficult.

### **2.6.4 Mixed methods that combine edge and region based methods**

In the year 2000, researchers such as Jianbo Shi and Jitendra Malik, came up with more robust hybrid methods that used combination of edge based and region based methods. Various methods such as Morphological watershed segmentation, Variable surface fitting and graph based methods such as Normalized - cut, Min - cut, Max - cut based image segmentation became popular during this time. The segmentation method



used in this research falls under this category and is discussed in detail in the next section.

### **3. Materials and Methods**

Native varieties of pecan nuts were obtained during 2004-2005 harvesting season from research farms at Oklahoma State University and several commercial pecan growers in Oklahoma. Nuts were stored and preserved in cold chambers maintained at 5 °C. The live pecan weevil larvae were obtained by directly breaking open the pecan nut. Some of the live weevil larvae were preserved in 95 % ethanol for future use. The diffuse reflectance measurements were performed on the dead, live and preserved weevil larvae.

One hundred nuts were randomly chosen from each lot of cultivars and numbered. Out of these, two groups of 50 randomly chosen nuts were formed. Group I was used for training and group II was used for testing and validation.

#### **3.1 Diffuse reflectance measurement**

VIS/NIR spectroscopy was used to measure relative reflectance or absorbance rather than absolute reflectance or absorbance. This measurement was referenced to a white lambertian surface diffuse reflectance measurement standard. The white Spectralon® panel from Labsphere, North Sutton, NH, had a correction factor of 0.99 in the UV/VIS/NIR spectral region. However, it was assumed to be 1.0 for simplification. Reflectance can either be expressed as a percentage (%R) relative to the reflection from a standard surface or as a factor between 0 and 1.

The reflectance measurement (% R) is given by:

$$\% R_{\lambda} = \left( \frac{S_{\lambda} - D_{\lambda}}{R_{\lambda} - D_{\lambda}} \right) * 100\% \quad (3.1)$$

Where,

$S_{\lambda}$  = Sample intensity of each wavelength,

$D_{\lambda}$  = Dark intensity of each wavelength

$R_{\lambda}$  = Reference intensity of each wavelength

$\lambda$  = at each wavelength

The reflectance data can be linearized by taking log of 1/R (Hruschka, 2001) as shown in the equation below:

$$A_r = \log \frac{1}{R} \quad (3.2)$$

Where  $A_r$  = apparent absorbance or optical density in reflectance mode

As the sample and reflectance standard were satisfactorily thick enough and opaque, none of the incident radiation passed through it. So the transmitted radiation  $I_t =$

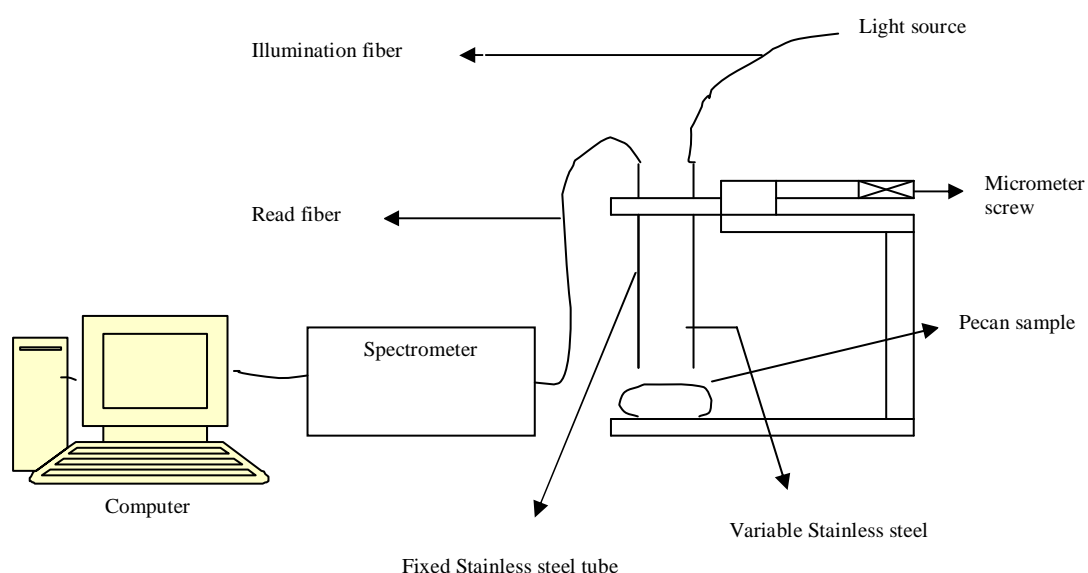
0. Equation (2.1) can now be written as:

$$I_o = I_r + I_a \quad (3.3)$$

So, now the equation (3.2) becomes:

$$A_r = \log \frac{1}{R} = \log \frac{I_o}{I_r} = \log(I_o) - \log(I_r) \quad (3.4)$$

The absorbance or optical density is a dimensionless number and has no units. The positioning of the sample and reference relative to the incident light source has significant effect on the  $A_r$  measurement (Roberts et al., 2004). Therefore, it is mandatory that the optical light source, sample and reference standard be absolutely stable. Any dirt, moisture or oil on the reference standard would result in the measurement of erroneous absorptions. Therefore, it is also required that the reference standard be clean. A typical reflectance experiment configuration is shown in figure 3.1.



**Figure 3.1 A typical configuration for a reflection experiment**

The VIS/NIR Labspec Pro spectrometer from Analytical Spectral Devices Inc., Boulder, CO, was used to obtain the diffuse reflectance measurements of the pecan weevil larvae and nutmeat. Specifications for the spectrometer are listed in table 3.1.

**Table 3.1 Specifications of Labspec Pro spectrometer**

Spectral Range	350-2500 nm
Spectral Resolution	3 nm @ 700 nm 10 nm @ 1400/2100 nm
Sampling Interval	1.4 nm @ 350 - 1050 nm 2 nm @ 1000 - 2500 nm
Detectors	One 512 element Si photodiode array 350 - 1000 nm Two TE cooled, InGaAs photodiodes 1000 - 2500 nm
Scanning Time	100 milliseconds
Photometric Linearity	Less than 1%
Wavelength Accuracy	$\pm 0.8$ nm
Wavelength Repeatability	$\pm 0.02$ nm

### 3.2 Analyzing spectral measurement data

The spectral location (*nanometers*), amplitude (*intensity*) and the width (*bandwidth*) of the peak comprehensively define the physical and chemical structure of the sample. It is therefore necessary that the spectral signal has minimum noise and have sufficient strength (Roberts et al., 2004).

The signal-to-noise ratio (SNR) of the reflectance data is influenced by the integration time, dynamic average of samples, average of adjacent pixels and dark current. The influence of noise increases with the *integration time*. The SNR improves by the square root of number of scans averaged. The advantage of higher SNR has to be traded with the amount of time required to scan. An integration time of 100 ms was used

in this research. Each spectrum was collected and the running average was calculated. Ten sample scans were averaged for this application. The number of *adjacent pixel elements* to be averaged was set to zero (no averaging) as higher values lead to loss of spectral resolution. This averaging was applied after determining the potential wavebands. As the loss of spectral resolution is higher in non-weighted averaging techniques, a weighted *Savitzky- Golay* averaging was performed after determining the spectral wavebands.

Dark current introduces variation in noise behavior as a function of integration time and operating temperature (McFee, 2005). Therefore, the dark measurements were conducted at the same integration time that was used for the spectrum measurement with the light source switched off. The difference in the value between reference level and data level ranges from 20 to 70 mV. As a consequence, the uncorrected dark current reduces the system dynamic range (Barn, 2005). In order to measure the dark current, same operating modes and environmental conditions were used. The entire data set was corrected by subtraction of the averaged dark current on a pixel-by-pixel basis.

### **3.2.1 Principal component analysis (PCA) and derivative analysis**

Principal component analysis is a data reduction technique through which the information from different wavebands is condensed into a few principal components. It is used to reduce the dimensionality of the data while minimizing loss of information. In this research PCA was used to capture the maximum variance between pecan weevil larvae reflectance and nutmeat reflectance.

PCA analysis involved five basic steps. The first step was to *mean center* the data. That was to construct the data with zero mean. This was done by subtracting the

average from the original data set across each dimension. The second step involved calculating the **covariance matrix**. The covariance matrix defines the variability in the data set. The covariance  $C_{x,y}$  between the two dimensions (X and Y) in the data set was calculated by following equation:

$$\text{cov}(X, Y) = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{n-1} \quad (3.5)$$

The resulting  $C_{x,y}$  value would be positive if X and Y tend to increase together, negative if they tend to decrease together and 0 if they are independent. The third step involved calculating the **eigenvectors and eigenvalues** of the covariance matrix. The eigenvalue gives us the magnitude of variance and eigenvector associated with that eigenvalue gives us the direction of the variance. This eigenvectors of the covariance matrix are orthogonal to each other. Next step of PCA was forming the **feature vector**. Once the eigenvectors were found from the covariance matrix, the next step was to order them by looking at their eigenvalues. The eigenvector with the largest eigenvalue is the first principle component (**PC1**). The eigenvector with second largest eigenvalue is the **PC2** and so on. This defines the order of significance. Taking the eigenvectors of importance and forming a matrix with these eigenvectors, as columns of the matrix would construct the feature vector.

$$\text{Feature vector} = (\text{eig1, eig2, eig3} \dots \dots \dots) \quad (3.6)$$

The last step in PCA was **deriving the new data set**. This step was the most important step in PCA. In this step matrix multiplication of the feature vector with the original mean centered data set was performed to get the new data set. This new data set was expressed with **PC1** as its axis.

Derivative analysis uses the derivatives of a spectrum to characterize the spectral data. The original spectrum may have hundreds of data values, therefore by extracting important features from the data, it can be represented less tediously. The derivative of a spectrum also improves the visual interpretation of the data.

The *finite difference* method is the simplest method that estimates the derivative by calculating the difference between two adjacent points. This method is not generally used as it enhances noise in the spectrum. Therefore, the *Fourier transform* was used to calculate the derivative in the frequency domain. It uses the information in the neighborhood (data points on either side) of a particular wavelength to calculate the derivative rather than the difference between adjacent data points. This avoids the problem of noise enhancement. The size of the neighborhood is an important factor. If the size of the neighborhood is too small it would approximate finite difference method and if it is too large it would smooth out the important information. Thus the size of the neighborhood becomes a critical factor when using the *Fourier transform*.

### **3.3 Selection of narrow band-pass interference filter and NIR enhanced camera**

Narrow band pass interference filters are widely used in food inspection and processing lines (Williams and Norris, 2001). The quality of a narrow band pass interference filter depends on various factors such as peak transmittance at central wavelength, FWHM (full-width half- maximum) and upper and lower cut-off frequencies. According to William and Norris, the peak transmittance at the central wavelength should be at least 40 % and the filter should have FWHM between 10 nm to



20 nm for it to be suitable as a narrow band-pass interference filter for NIR instrumentation.

The interference filters are designed to operate in a specific temperature range. The operating temperature for a typical filter ranges from  $-40^{\circ}\text{C}$  to  $70^{\circ}\text{C}$  and the maximum allowable change in temperature is  $5\text{-}10^{\circ}\text{C}$  per minute. The central wavelength shifts linearly in the direction of the temperature change. This change is usually from  $0.01$  to  $0.03\text{ nm per }^{\circ}\text{C}$ . Operation beyond the specified temperature range may result in irreversible shift in central wavelength (Oriel, 2005). The peak transmittance and the bandwidth may also change with temperature. A slow drift in central wavelength may result due to aging and high humidity. Good quality filters are hermetically sealed to prevent negative effects of aging.

Various hardware instruments such as *filter wheel*, *adapters* or *step-up*, *step-down rings* can be used to mount and hold filters. The filter can directly be mounted on the camera lens or it can be held discretely from the camera and lens unit. Two different types of settings are possible in principle. The filter can either be mounted on front of the lens or rear end of the lens.

Mounting the filter in front of the lens poses several problems. The filters are coated to reduce the flare, resulting in the reduction of light that passes through the lens. The light also requires passing through two additional filter surfaces prior to entering the lens that can degrade the quality of an image. Therefore, the filter was mounted at the rear end of the lens

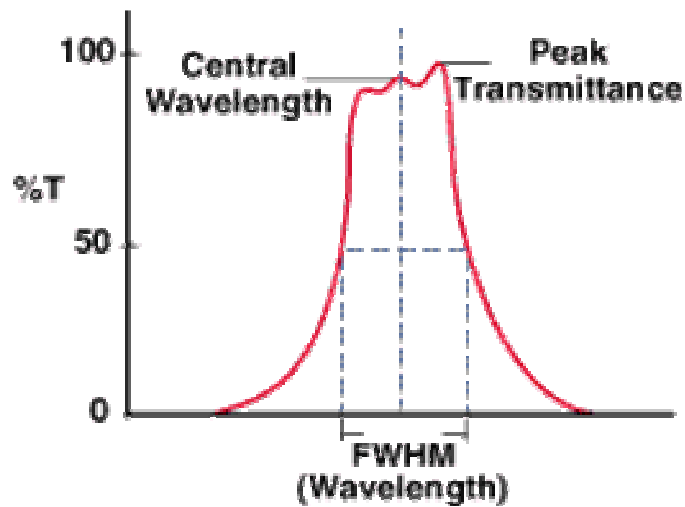
The major disadvantage of mounting the filter onto the lens is that only one filter can be mounted at a time. Some applications may require the use of a single camera with

multiple filters. In such a situation, filter wheel is a suitable choice. The filter wheel allows us to mount 6 to 12 filters. Computer controlled filter wheels are available that allow us to take images at different wavebands in a controlled manner without human intervention.

Four narrow band-pass interference filters between 850 nm and 980 nm were purchased from Edmund Optics, Barrington, NJ. The detailed specifications for the filters are listed in table 3.2. The characteristic of the Edmund optics filters is shown in figure 3.2.

**Table 3.2 Specifications of narrow band-pass interference filter**

Center Wavelength Tolerance	$\pm 3.0\text{nm}$
FWHM Tolerance	855nm, 905nm, 940nm and 980nm: $\pm 2.5\text{nm}$
Min. Transmission at CWL	40% to 50%
Blocking	$\leq 0.005\%$ from 200 to 3500nm (excluding band pass)
Clear Aperture	25mm Dia: $0.825'' \pm 0.005''$
Mount Thickness	5.5mm nominal
Surface Quality	80-50
Operating Temperature	$-20^{\circ}\text{C}$ to $75^{\circ}\text{C}$ ( $-4^{\circ}\text{F}$ to $167^{\circ}\text{F}$ )

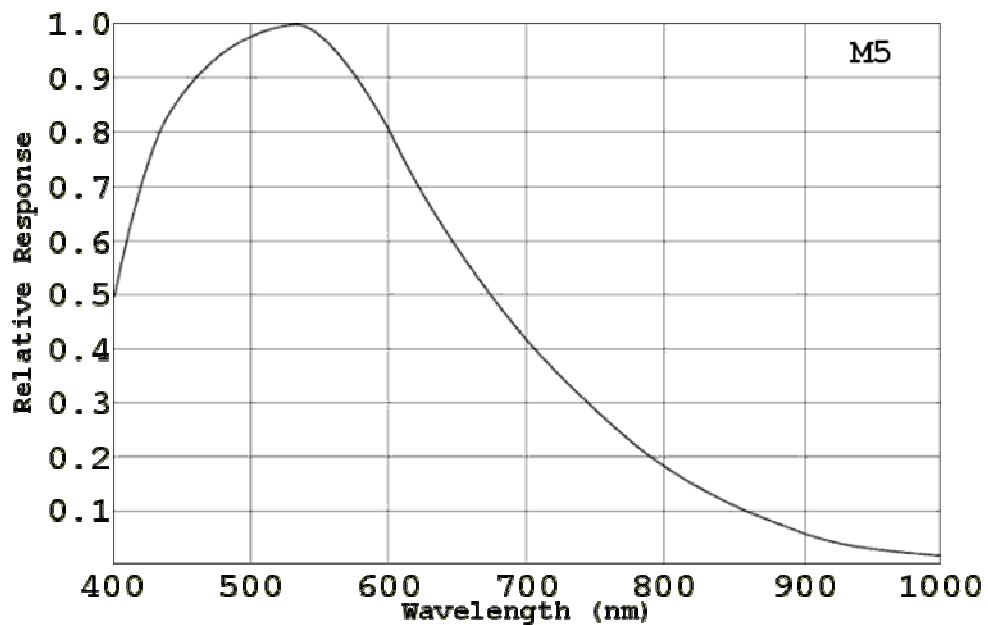


**Figure 3.2 Narrow band pass interference filter from Edmund optics**

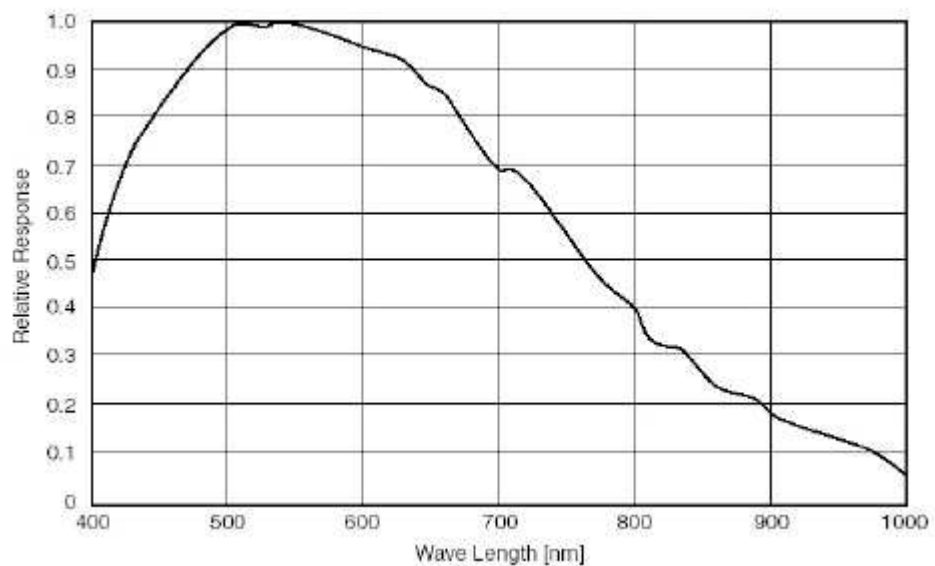
The images were taken with a NIR enhanced camera, AVT Dolphin F 145-B from Allied Vision Technologies, Newburyport, MA. It is a monochrome IEEE 1394 SXGA+ C-Mount camera, equipped with a mega pixel SONY 2/3" type progressive CCD-array. Specifications for the camera are listed in table 3.3. The spectral response of the AVT Dolphin camera is shown in figure 3.4. Figure 3.3 shows the spectral response of a typical silicon CCD. The enhanced silicon CCD of AVT Dolphin camera shows about 20 percent spectral sensitivity at around 900 nm as compared to 5 percent sensitivity of an ordinary silicon CCD. The higher spectral sensitivity is of critical importance as the images were acquired in NIR waveband between 850nm to 980nm.

**Table 3.3 Specifications of AVT Dolphin camera**

Image Device	2/3 Type progressive scan b/w SONY Silicon CCD
Resolution	1392 (H) x 1040 (V)
Lens Mount	C-Mount
ADC	12 Bit
Data Path	8 Bit or 10 Bit
Frame Rates	1,875 fps - 15 fps; External Trigger
Gain Control	Manual 0 – 30 dB (0,035 dB/step)
Shutter Speed	Time base 1, 2, 5, 20, 50, 100, 200, 500, 1000 $\mu$ s
Transfer Rate	400 Mb/s
Digital Interface	IEEE 1394; DCAM V1.3



**Figure 3.3 Spectral response of a typical silicon CCD**

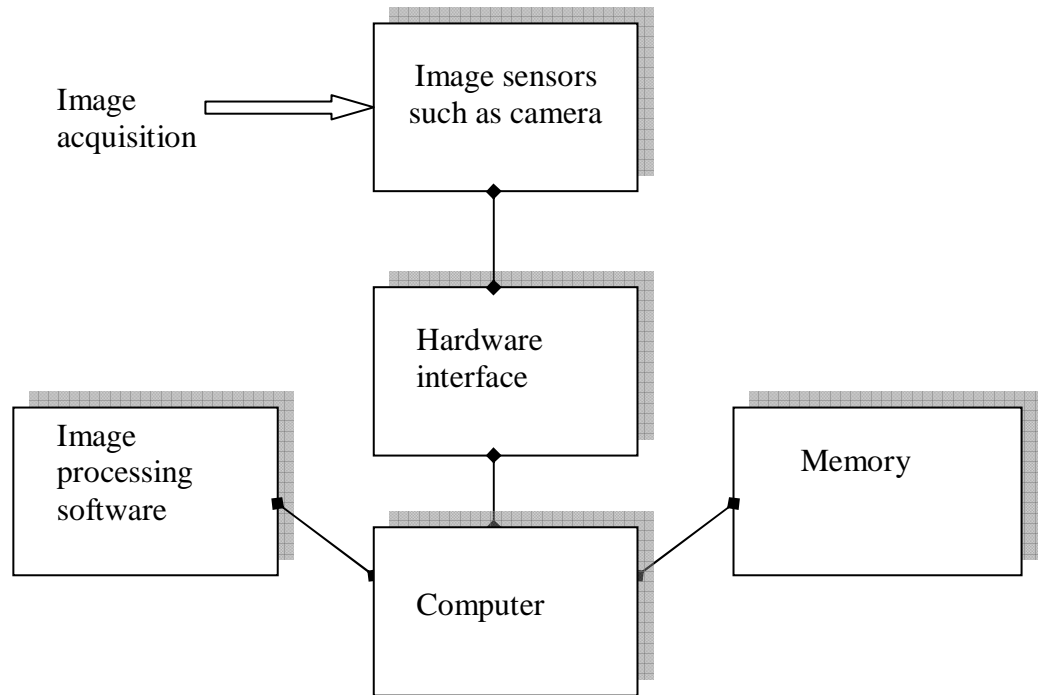


**Figure 3.4 Spectral response of the AVT Dolphin 145-B camera**

### **3.4 Acquiring and processing an image**

#### **3.4.1 Components of an image processing system**

The most important component of an image processing system is the computer. The Pentium(R) 4 CPU, 512 MB RAM, with processing speed of 2.8 GHz was used for acquiring and processing images. Another important component is the image processing software. MATLAB was used in processing the acquired images. Typical diagram of an image processing system is shown in figure 3.5.



**Figure 3.5 Components of image processing system**

IR Lamp (EKE-IR) from Illumination Technologies, Inc., East Syracuse, NY, was used as the lighting source. The output of this lamp was enhanced and limited within the 800nm to 1.5 $\mu$ m spectral range. Due to the lower sensitivity of silicon CCD in NIR waveband (fig 3.4), the images were acquired with a high gain factor of 200. The noise in the image increases with the gain. As the addition of randomly distributed noise variables is a Gaussian function, the noise was assumed to have a Gaussian distribution. The histogram analysis of the acquired image revealed no shift in the DC component and noise with a unit standard deviation. Therefore, a standard Gaussian low pass filter with a zero mean and unit variance was used to eliminate the noise.

### **3.5 Image Segmentation**

The acquired images were segmented by using two methods. Two different approaches of segmentation was used to aid comparison of traditional method of masking

followed by morphological processing to a novel active contour based snake energy minimization algorithm. The detailed description of the two methods is presented below.

### **3.5.1 Method 1**

The first method was a popular approach where the Gaussian filtered images are thresholded and processed by binary morphological operators. Fifty training images and fifty testing were acquired under constant illumination conditions, shutter speed and f-stop. A Gaussian low pass filter with zero mean and unit variance was used to restore the image. A threshold was determined for each of the training images in the database individually. The images were then processed by using morphological cleaning followed by morphological closing with a disk of size 3.

The morphological cleaning operation involves removing the isolated pixels (1's surrounded by 0's). The morphological closing can be defined as dilation followed by erosion. An average threshold value for 50 training images was determined and applied to 50 test images. The Gaussian filtered images were also segmented by Canny edge detector. The Canny edge detection method is explained below:

#### **3.5.1.1 Canny edge detector**

Canny edge detector is often referred to as an optimal edge detector. The several criteria considered by Mr. Canny to optimize the edge detection algorithm are as follows:

1. ***Signal to noise ratio:*** The filter should respond more strongly to the edge at  $x = 0$  than to noise.
2. ***Edge Localization:*** The filter response should reach a maximum close to  $x = 0$ . That means the filter should give maximum response at the edge than to noise.

3. **Low false positives:** There should be only one maximum of the response in a reasonable neighborhood  $x=0$ . That means that the filter should have signal response to single edge unlike LoG where the filter has multiple responses to a single edge.

Keeping in mind the above three criteria Canny derived an edge detector that satisfies all the above three criteria to the maximum. The first step of the Canny edge detector is to take the first derivative of the Gaussian:

The first order derivative along x and y direction are calculated separately. The first order derivative along x – direction is given by:

$$f_x = \frac{\partial f}{\partial x} = K_{\nabla_x} ** (G_\sigma ** I) = (\nabla_x G_\sigma) ** I \quad (3.7)$$

Where,

$$\nabla_x G_\sigma = \frac{-x}{2\pi\sigma^4} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \quad (3.8)$$

The first order derivative along y – direction is given by:

$$f_y = \frac{\partial f}{\partial y} = K_{\nabla_y} ** (G_\sigma ** I) = (\nabla_y G_\sigma) ** I. \quad (3.9)$$

Where,

$$\nabla_y G_\sigma = \frac{-y}{2\pi\sigma^4} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \quad (3.10)$$

After computing the gradient magnitude along the two directions, the combined gradient magnitude is written as:

$$|\nabla f(x, y)| = \sqrt{f_x^2 + f_y^2} = \text{rate of change of } f(x, y) \quad (3.11)$$

The direction of the gradient response is given by:



$$\angle \nabla f(x, y) = \tan^{-1} \left( \frac{f_y}{f_x} \right) = \text{orientation of change of } f(x, y) \quad (3.12)$$

The gradient magnitude image is thresholded to get the points that are potentially the edge points. After thresholding the gradient magnitude image, we perform non-maximal suppression. Non-maximal suppression is the process of thinning where a potential edge pixel is checked for local maxima along the gradient direction. The gradient direction is perpendicular to the edge direction.

### 3.5.2 Method 2

In the second method, an active contour based snake energy minimization algorithm (Ying et al., 2005) was used to segment out the weevil larvae from nutmeat. This method is a combination of edge-based method and seeded region growing (SRG). Edge detection was performed by using gray scale morphological dilation residue edge detector. Before the dilation residue edge detector is applied, the images were first closed to remove the noise after which the contrast enhancement was performed by using top hat transformation.

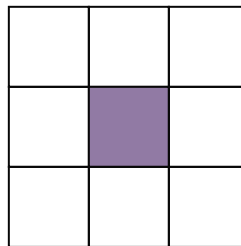
The seeds points were deployed in the edge-detected image and region growing was performed by using an active contour based snake energy minimization algorithm (Ying et al., 2005). The detailed description of the algorithm is now presented.

#### 3.5.2.1 Active contour based snake energy minimization algorithm

Grayscale morphology is simply a generalization of binary morphology to images with multiple bits per pixel, where the *Max* and *Min* operations are used in place of the OR and AND operations. Grayscale erosions and dilations produce results identical to the nonlinear minimum and maximum filters. The significance of nonlinear operations, as opposed to linear operations, is that *nonlinear operations can be used directly for making*

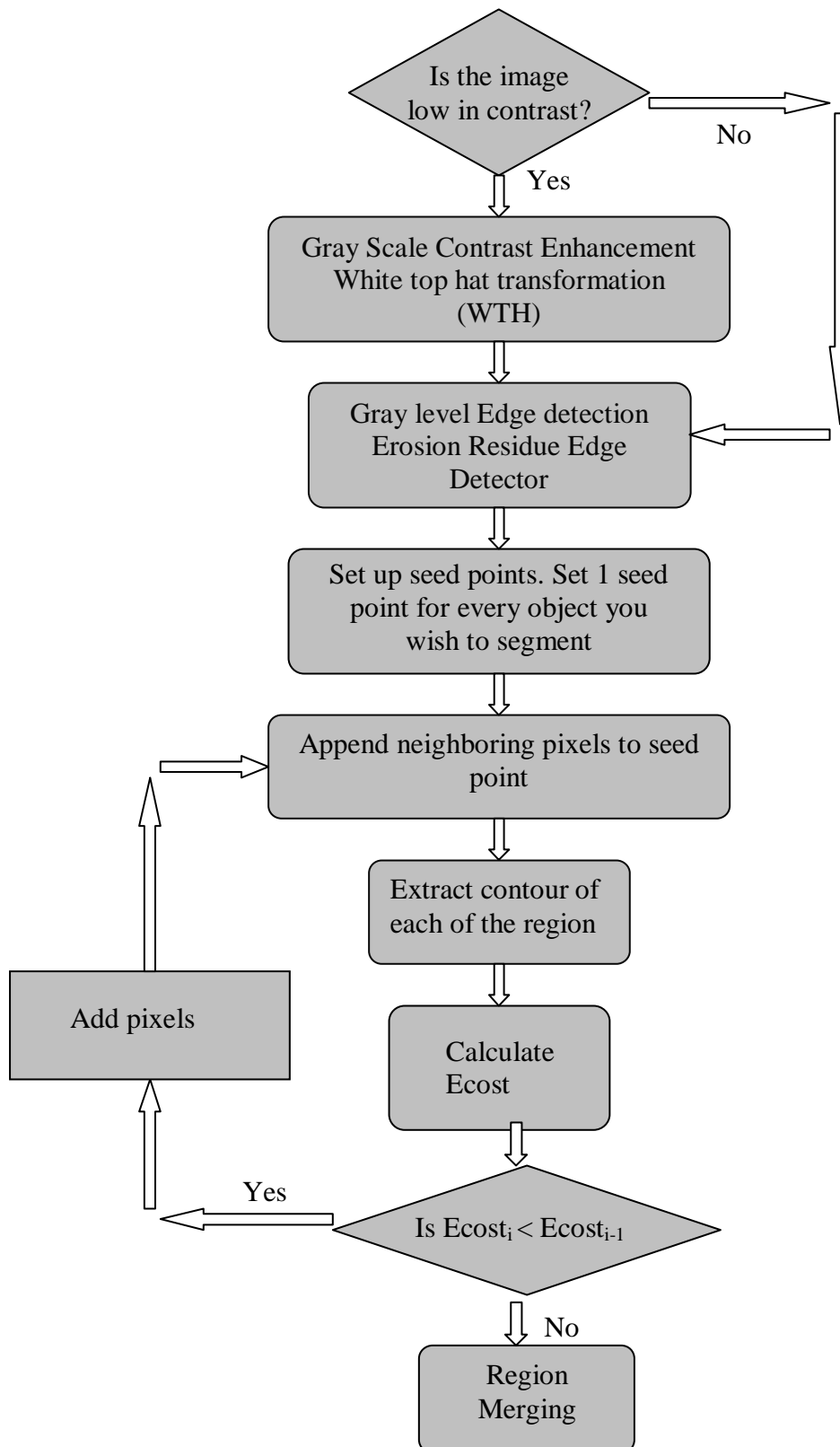
*decisions about regions of the image.* Nonlinear operations are therefore of particular use in *image analysis*. The structuring element used here is a 3D structuring element where the first two dimensions give the neighborhood information while the third dimension gives the gray scale magnitude.

The structuring element used for the project is shown in figure 3.6.



**Figure 3.6 Structuring element**

The figure 3.7 shows the flowchart of the complete algorithm that was used to segment out the pecan weevil larvae from the nutmeat using an active contour based snake energy function.



**Figure 3.7 Complete Algorithm Flowchart**

The gray scale morphological contrast enhancement operation was performed by the following enhancement operator:

$$K^{th} = F + WTH(F) - BTH(F) \quad (3.13)$$

Where

$$WTH(F) = F - (F \circ B) \text{ (White top hat transformation)} \quad (3.14)$$

$$BTH(F) = (F \bullet B) - F \text{ (Black top hat transformation)} \quad (3.15)$$

This combined contrast enhancement operator was applied iteratively to enhance the bright and dark regions of the image.

Gray scale morphological edge detection can be performed by using either erosion residue edge detector or dilation residue edge detector. The dilation residue edge detector was used in this research. Mathematically, the dilation residue edge detector is defined as:

$$E_d(F) = (F \oplus B) - F \quad (3.16)$$

### 3.5.2.2 Region growing

The goal of region growing is to divide the domain  $R$  of the image  $F$  into sub regions  $\{R_i, i = 1, 2, m\}$  so that

$$\begin{aligned} R &= \bigcup_{i=1}^m R_i \\ R_i \cap R_j &= 0, i \neq j \end{aligned} \quad (3.17)$$

### Seeded region growing (SRG)

Seeded region growing algorithm (SRG) is a new approach where instead of controlling region growing by fine-tuning homogeneity parameters, SRG is controlled by choosing a small number of pixels, known as seeds. These seed pixels are chosen by user according to his own opinion what should be regions to extract on the image. The basic

idea of region growing is to have a set of seed points and append to it the neighboring pixels that have got property similar to seed points. The region is grown in such a way that the homogeneity is preserved in each region  $R_i$ .

Firstly, the seeds were grouped into  $n$  sets:  $C_1, C_2 \dots C_n$ . At each step of the algorithm one pixel was added to some of the sets  $C_i, i=1 \dots n$ . Consider, the state of sets  $C_i$ , after  $n$  steps. Let  $Y$  be the set of unallocated pixels which borders at least one of the regions, which means that  $Y$  is the set of all pixels which are on the borders of the regions that are formed until now. The set  $Y$  can be described as:

$$Y = \left[ x \notin \bigcup_{i=1}^n C_i \mid M(x) \cap \bigcup_{i=1}^n C_i \neq \emptyset \right] \quad (3.18)$$

Where  $M(x)$  is the set of immediate neighbors of the pixel  $x$ . For this research, 8-connectivity i.e. 8 neighbors surrounding each pixel is implemented.

### 3.5.2.3 Problem with traditional seeded region growing (SRG) and solution

For this hybrid method of image segmentation, the edge based information was combined with the seeded region growing algorithm. The region growing process was performed on edge extracted image. The edges extracted were not always continuous. The morphological dilation residue edge detector do produce edges that are relatively continuous than Canny edge detector, however still some discontinuity exist in the edge extracted image. This results in the traditional region growing algorithm to grow through the gap of the edge in an uncontrolled fashion.

To overcome this problem a constraint similar to snake energy function was introduced. The boundary condition is different from the snake energy approach because

the first one controls the growing process of each region and the last one controls the reaching process of an edge vector (Ying et al., 2005).

The idea of this function is to grow region in such a way as to minimize the energy function. The Cost energy function is given as:

$$E_{\text{cost}} = \sum_{i=1}^N (E_e(i) + E_r(i)) \quad (3.19)$$

Where,

$E_e(i)$  is the edge shape cost

$E_r(i)$  is the reached edge cost

The edge shape cost and reached edge cost can be defined as:

$$E_e(i) = \alpha_i |s_i - s_{i-1}|^2 + \beta_i |s_{i-1} - 2s_i + s_{i+1}|^2 \quad (3.20)$$

$$E_r(i) = -|255 - F(s_i)|/255 \quad (3.21)$$

$F(s_i)$  : Pixel value of the edge extracted image at position  $s_i$

The equation (3.20) defines the edge shape. The parameter alpha defines the spacing between the control points i.e. the smoothness of the contour and the parameter beta defines the curvature that can be allowed for the reached edge. The equation (3.21) takes into the consideration the gray level information of the region. The region growing process stops when it reaches the edge boundaries. The region growing process can be controlled by setting up the parameters alpha and beta. In equation (3.19), N represents the number of boundary point and the order of the boundary points is counterclockwise along the contour of the region.  $s_i$  is the coordinate of the ith boundary.

By minimizing the boundary condition  $E_{\text{cost}}$ , the region is attracted to grow to the boundary with the lowest cost, such as edge lines.

### 3.5.2.4 Region merging

Region merging is performed to combine two regions that share relatively small boundary length. Region merging is usually performed to automate the image segmentation process. Here, the seed points are randomly set up and the regions that have relatively small boundary length between them are merged. One of the ways to perform the region merging is to set a threshold and merge the regions where the number of edge points in the boundary falls below the threshold.

### 3.5.2.5 The problem faced with this algorithm

The contour points were taken in the counter clockwise direction along the contour of the region. The minimum distance criterion was used to find the next point in counter clockwise direction from the current point. This minimum distance criterion does not always work. For example consider the figure 3.8 shown below:

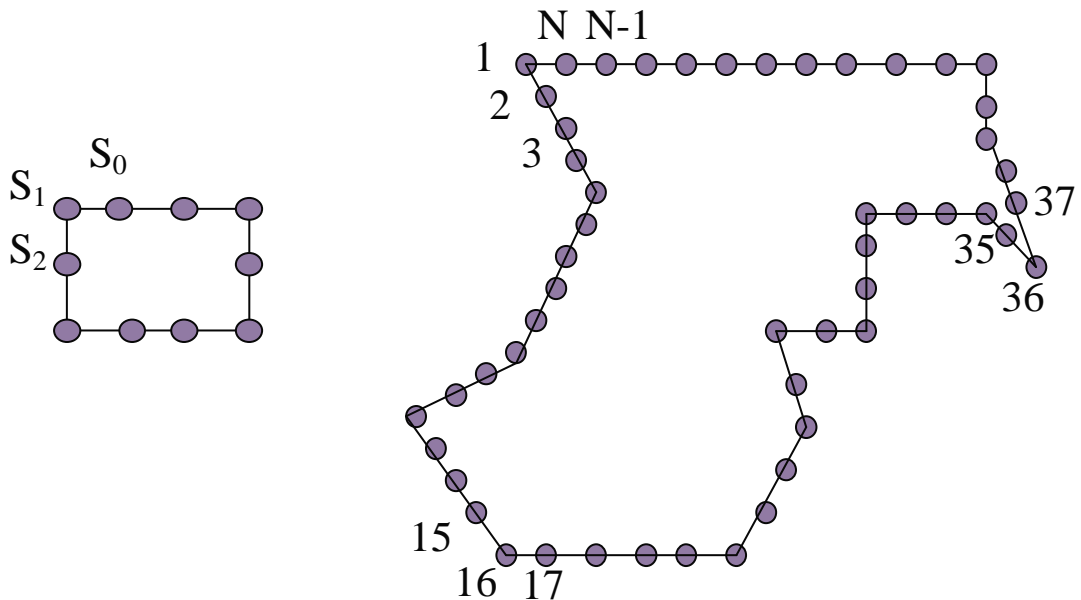


Figure 3.8a Rectangular boundary

Figure 3.8b Random boundary

The figure 3.8a shows the contour extracted from the rectangular region grown. S1 was chosen as the starting point. Then, the previous point S0 and next point S2 were determined. Proceeding in the counter clockwise direction, S2 was assigned as the current point and S1 as the previous point. The next point was chosen such that it was at the shortest distance from S2 in the 8-neighborhood and which was previously unallocated. This algorithm worked well for the object with rectangular boundaries.

However the region can grow in any shape as dictated by the edges. Consider the figure 3.8b where the region has grown to a shape as shown. Here for the pixel 34, 37 is the closest point instead of 35. So the algorithm would assign 37 as its next point missing out pixels 35 and 36. These unaccounted pixels may add up as the region grows up to a significant amount. This may give a wrong value for the energy function. It may create a problem in extreme cases and region may grow out of gaps and beyond control. However, for the images that were used, this problem was not encountered as the number pixels missed out are not high enough to change the energy value by a significant amount.



## 4. Results and discussions

### 4.1 Diffuse reflectance measurements

The figures 4.1, 4.2 and 4.3 show the diffuse reflectance measurements of pecan nutmeat and weevil larvae. The reflectance data was obtained for live weevil larvae, dead weevil larvae and weevil larvae preserved in 95% ethanol.

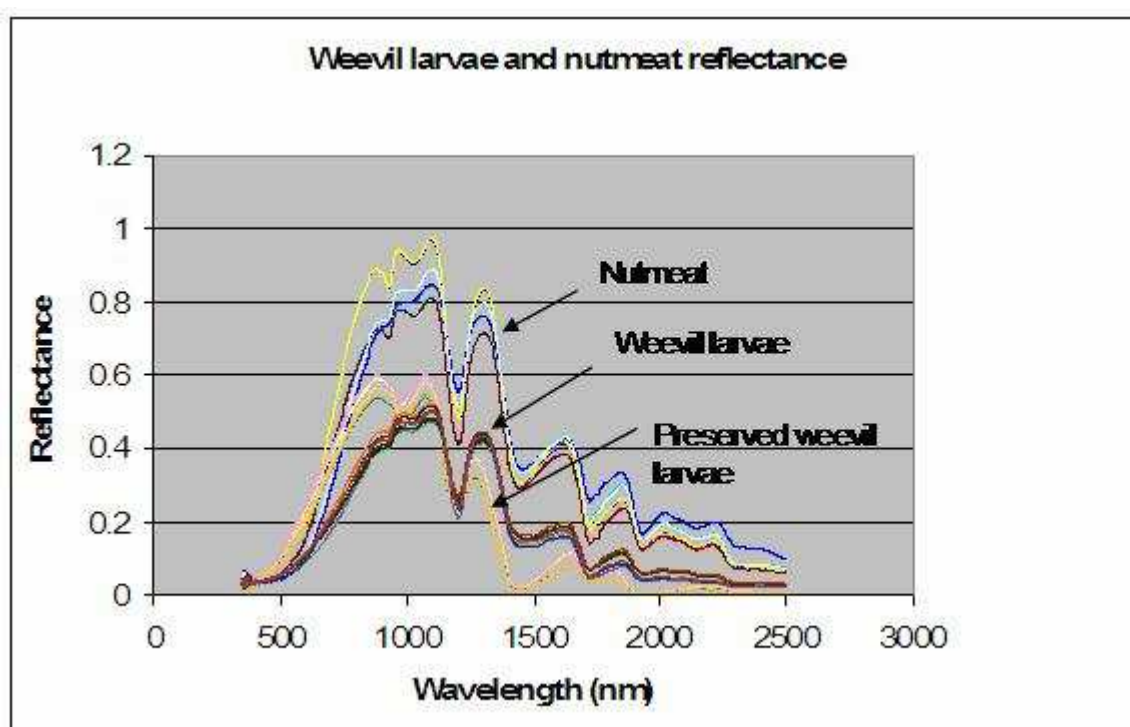


Figure 4.1 Diffuse reflectance measurements of pecan weevil larvae and nutmeat

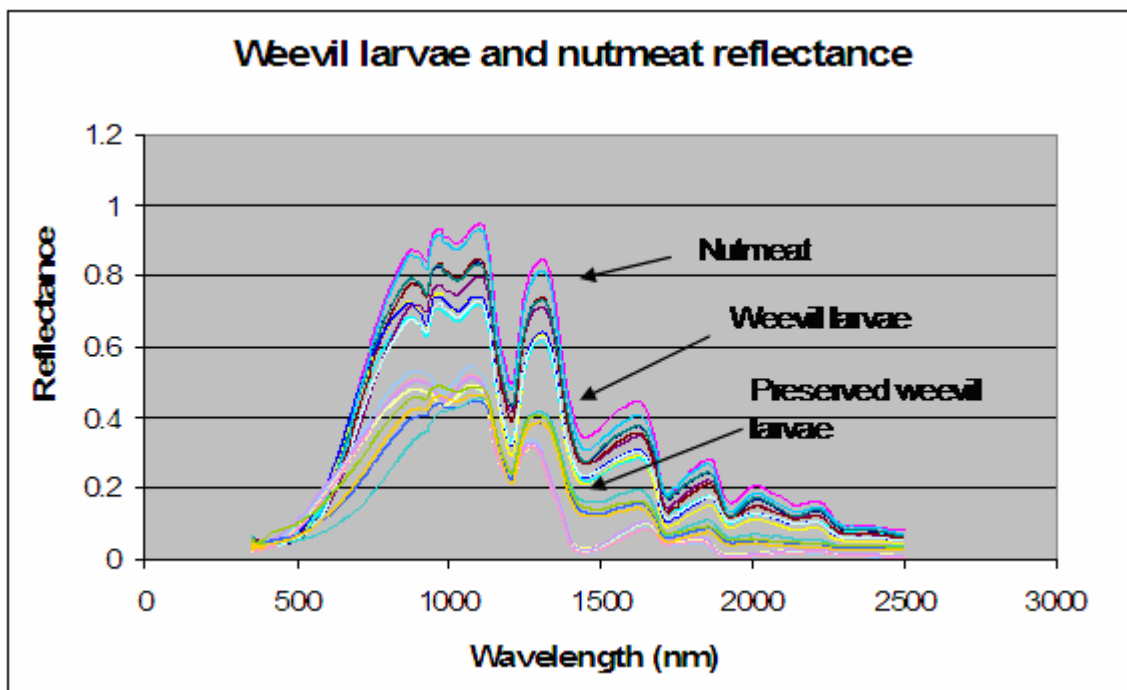


Figure 4.2 Diffuse reflectance measurements of pecan weevil larvae and nutmeat

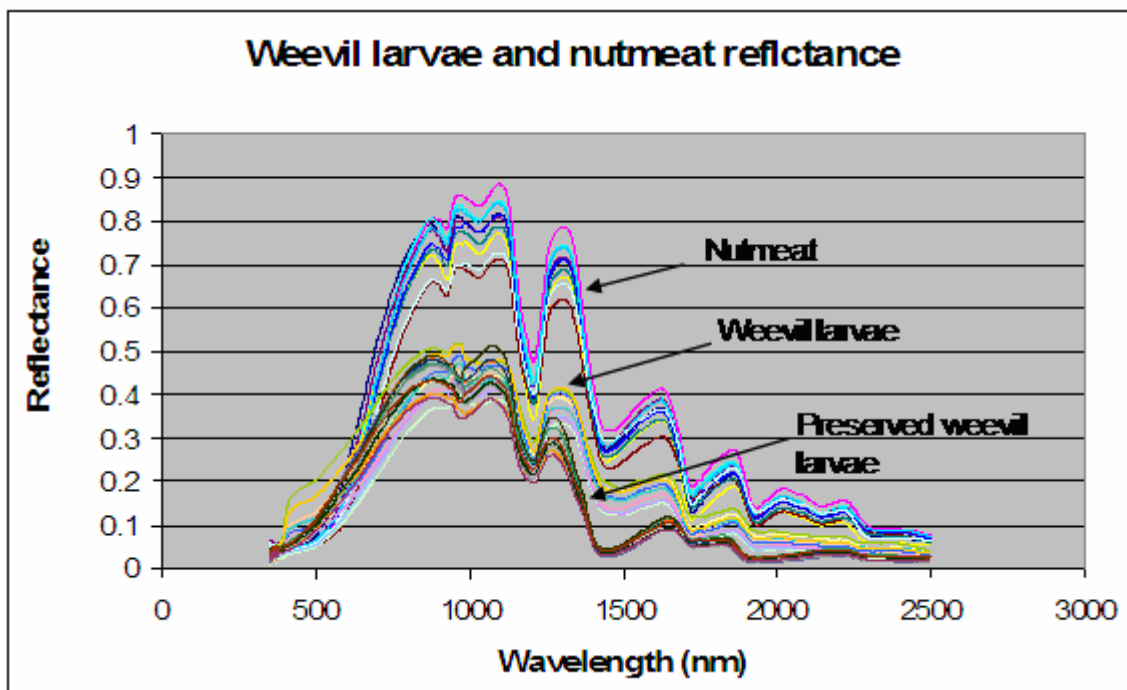
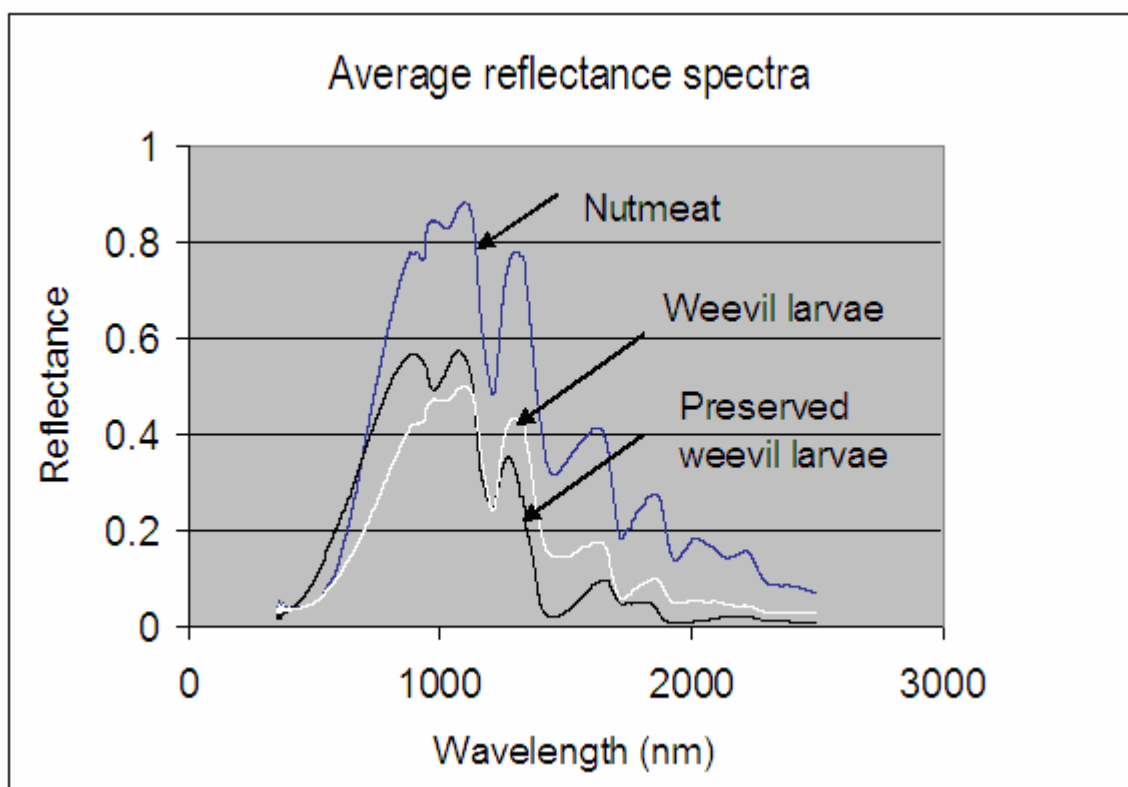


Figure 4.3 Diffuse reflectance measurement of pecan weevil larvae and nutmeat

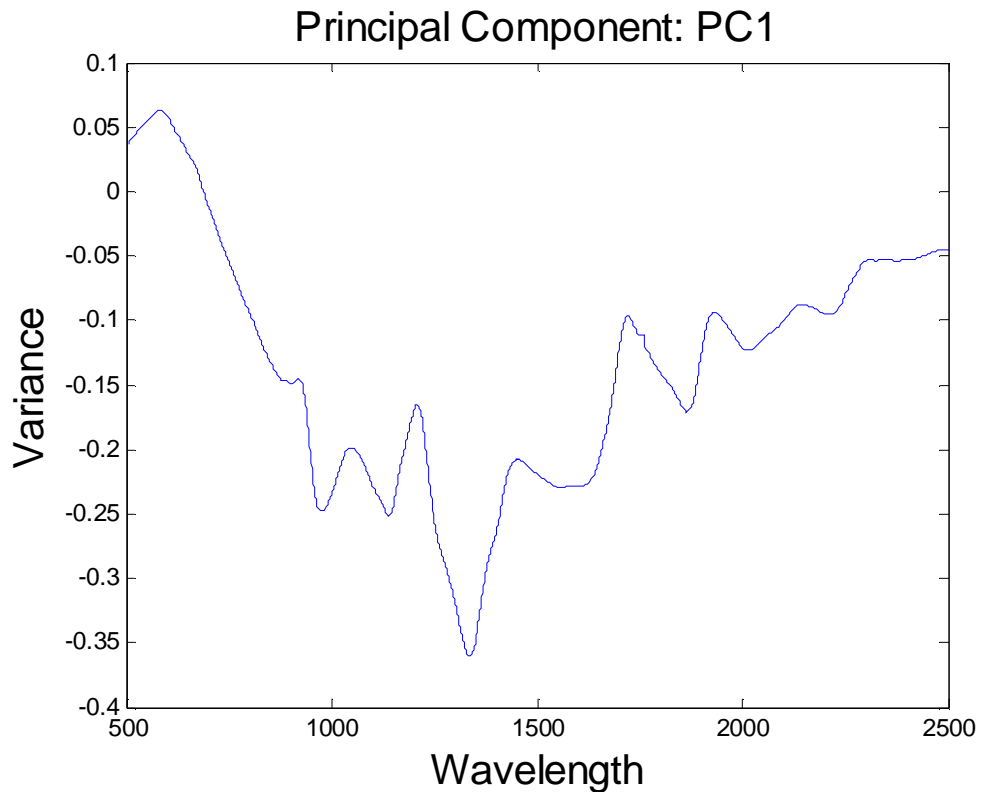
Visual analysis of figures 4.1, 4.2 and 4.3 show the higher reflectance of nutmeat. The average reflectance of weevil larvae (both dead and live) and nutmeat are shown in figure 4.4.



**Figure 4.4 Average Diffuse reflectance measurement of weevil larvae, weevil larvae preserved in 95 % ethanol and pecan nutmeat**

## 4.2 Data Analysis

The wavelength from 800 nm and 2500 nm showed a large variance between the pecan nutmeat and the weevil larvae. Principal component analysis was performed to capture the variance distribution between the weevil larvae and nutmeat. A single principal component (PC1) was sufficient to describe the variance. The figure 4.5 shows the result of PCA.



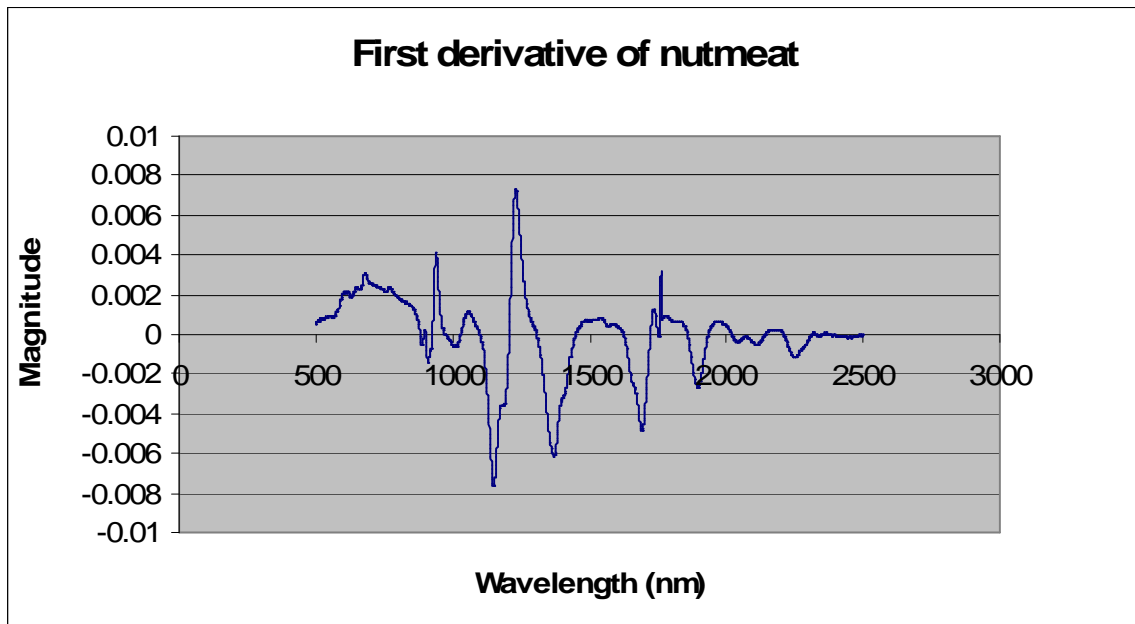
**Figure 4.5 Principal component analysis (PC1)**

Analyzing figure 4.5, the maximum variance was found to be at 1380 nm.

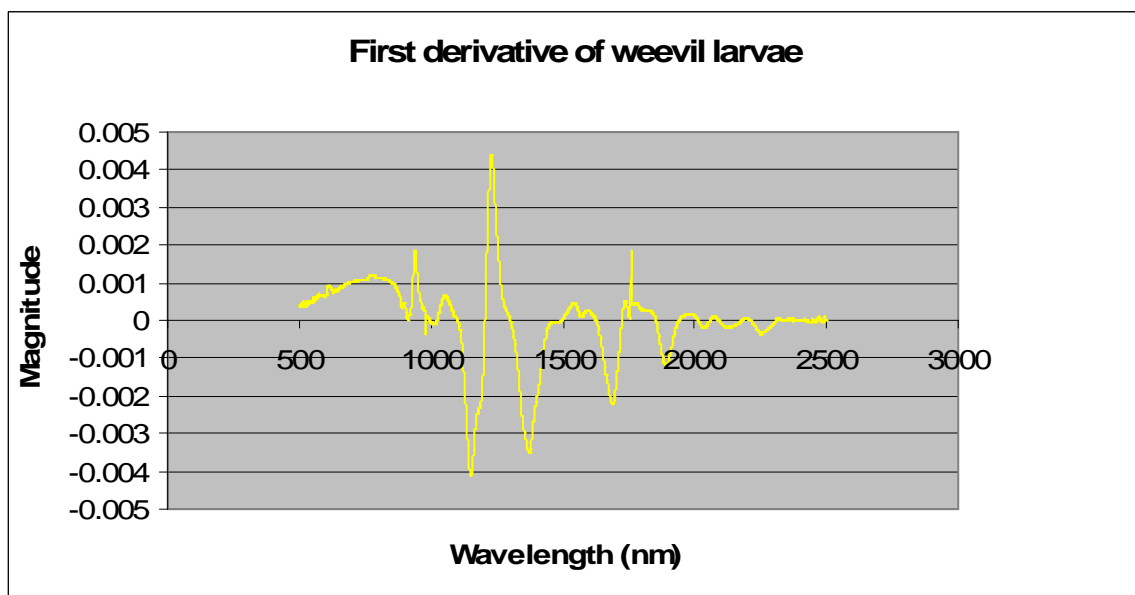
However, it was outside the spectral response range of a silicon CCD. The range 850 nm to 985 nm showed a considerable amount of variance with maximum variance at 982 nm. This wavelength was exploited to design a low cost multispectral imaging system.

The derivative analysis was used in conjunction with PCA. The point at which the derivative changes sign is of significant importance as it reveals the maximum variance between pecan weevil larvae and the nutmeat. The first derivatives graphs of the nutmeat, weevil larvae and preserved weevil larvae are shown in figure 4.6, 4.7 and 4.8 respectively. Figure 4.9 shows the derivative graphs superimposed for analysis. Three potential wavebands at 856 nm, 940 nm and 982 nm were identified for selection of narrow band pass interference filters. The images were acquired at 855 nm, 940 nm and

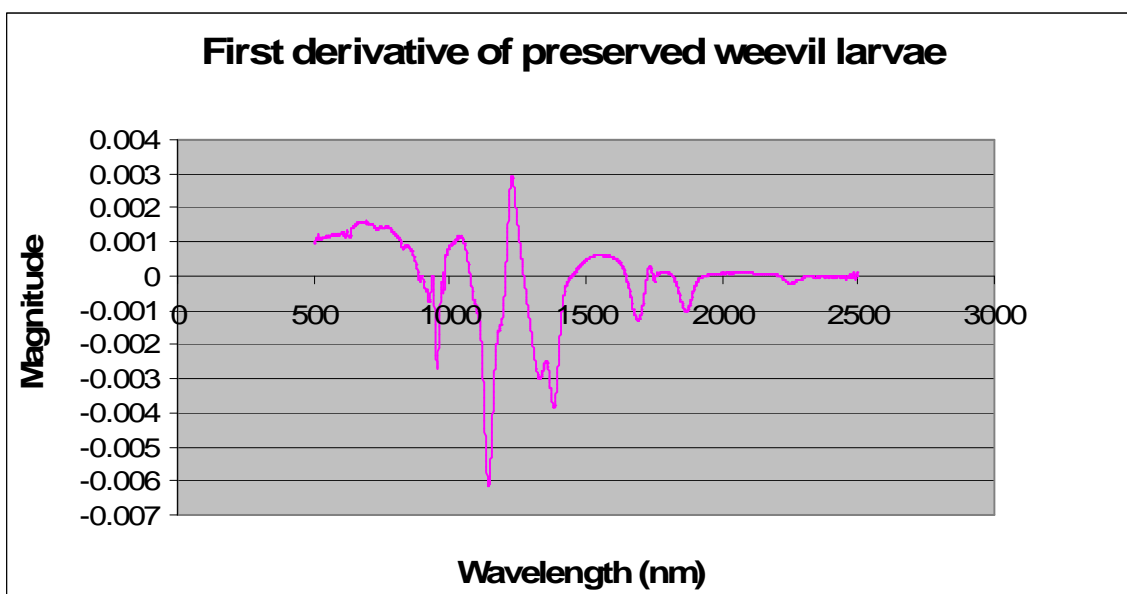
980 nm. However, only the images at 980 nm showed significant contrast difference between weevil larvae and nutmeat. This is due to the fact of higher variance at 980 nm than at 940 or 850 nm as revealed in figure 4.5.



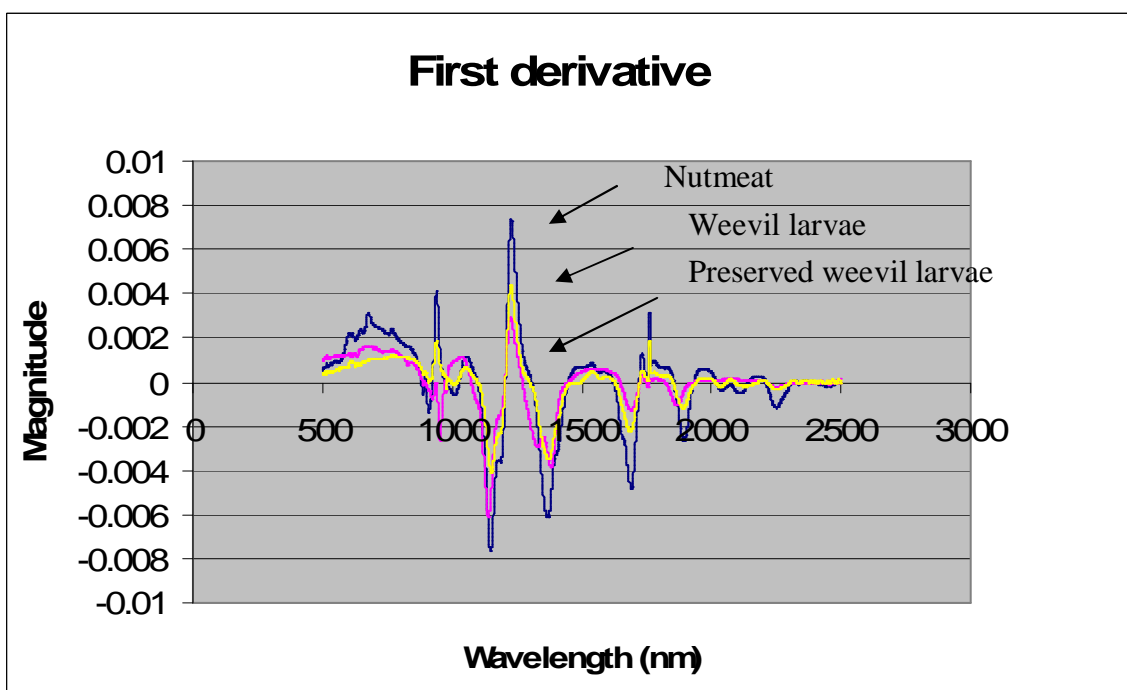
**Figure 4.6 First derivative of nutmeat**



**Figure 4.7 First derivative of weevil larvae**



**Figure 4.8 First derivative of preserved pecan weevil larvae**



**Figure 4.9 Derivative graphs superimposed**

By studying the results of the principal component analysis and the derivate analysis, four wavebands that captured the maximum variance between nutmeat and weevil larvae were determined. The wavelengths are given in order of maximum variance between pecan weevil larvae and pecan nutmeat.

1. 981 nm
2. 940 nm
3. 902 nm
4. 855 nm

### **4.3 Image acquisition and Canny edge detection**

The figure 4.10 shows an image acquired at 980 nm of the pecan nutmeat interspersed with weevil larvae before Gaussian filtering. The figure 4.11 shows the denoising result of the acquired image. The figure 4.12 shows the segmentation result using Canny edge detector. A variable threshold of 0.15-0.4 and a variable standard deviation of 5 – 41 were used to perform the Canny edge detection for each of the 50 images in the training database.

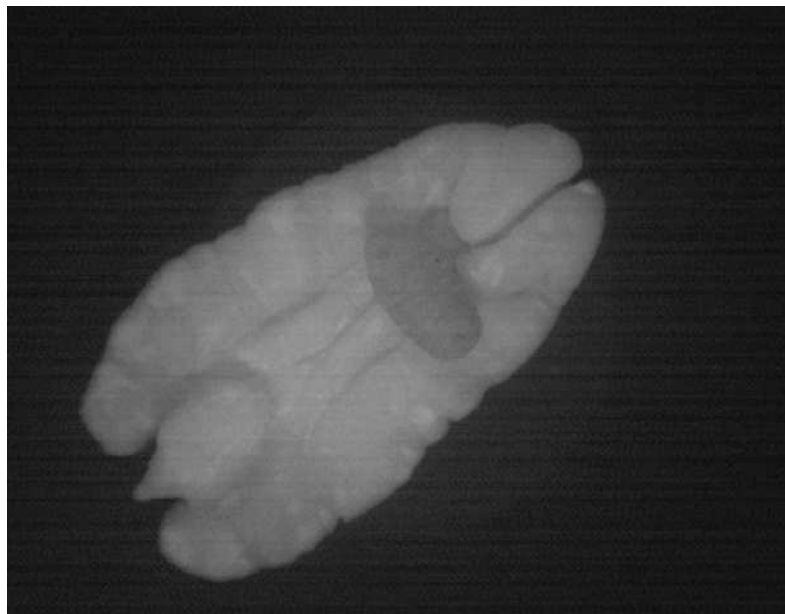
The section 4.4 shows the segmentation result for 50 training images by using method 1. In this method, thresholding was performed followed by morphological cleaning and closing with a disk as a structuring element. The images were acquired at 980 nm. These images were filtered by a Gaussian filter of zero mean and unit variance. The threshold value used for the 50 images in the training database is shown in figure 4.14. Average threshold was calculated to be 108. This threshold was used on the 50 test images with 74 percent correct classification. The type 2 error was determined to be 0 percent.

Image: 1



**Figure 4.10 Example image acquired at 980 nm**

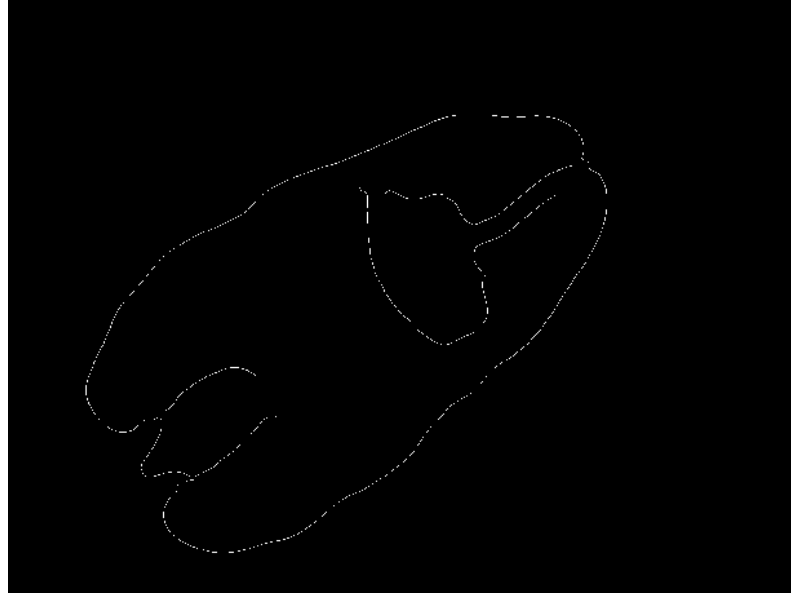
Noise reduction of Image: 1



**Figure 4.11 Gaussian filtered image**



Edge Detection of Image: 1

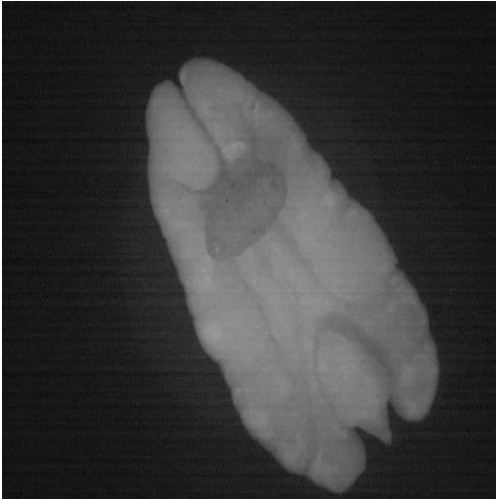


**Figure 4.12 Result of Canny edge detection**

#### **4.4 Segmentation results for training images using method 1**

The segmentation using Method I was performed on all the 50 images in the training database. The figure 4.13 illustrate the segmentation results obtained for five images at 980 nm. The classification accuracy for the 50 training images is shown in table 4.1.

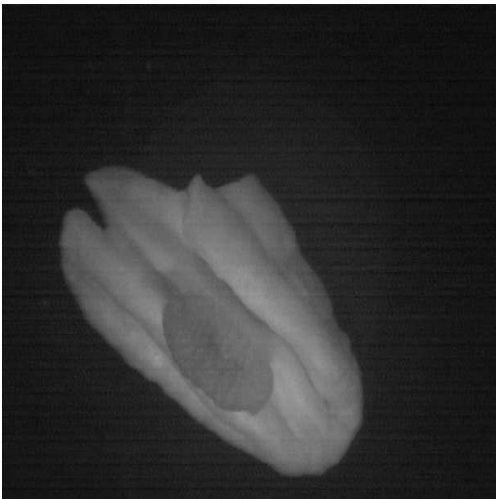
**Figure 4.13: Segmentation results for training images using method 1**



**Gaussian filtered image 1**



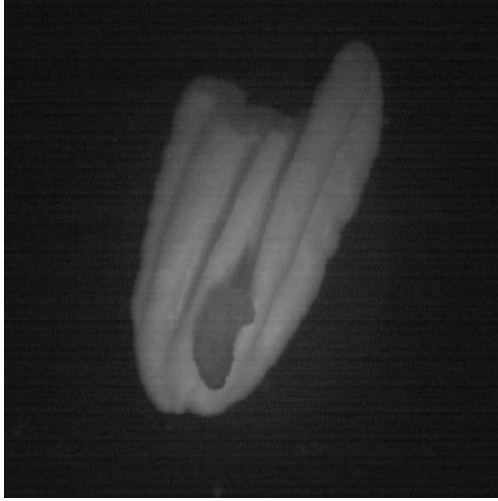
**Segmented image 1**



**Gaussian filtered image 2**



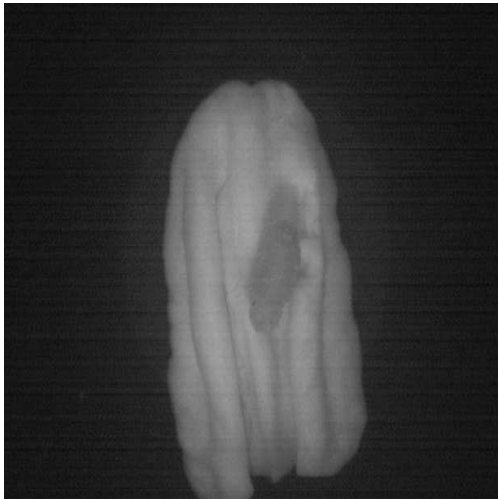
**Segmented image 2**



**Gaussian filtered image 3**



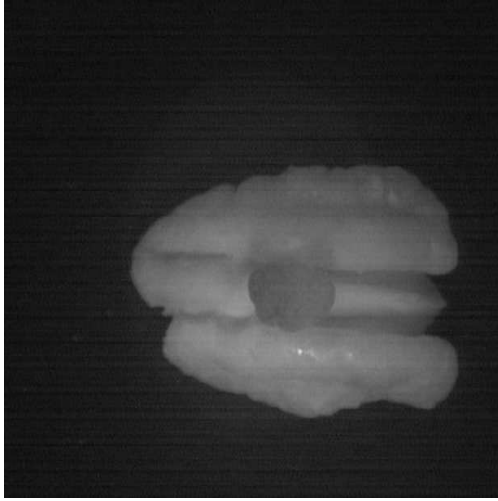
**Segmented image 3**



**Gaussian filtered image 4**



**Segmented image 4**

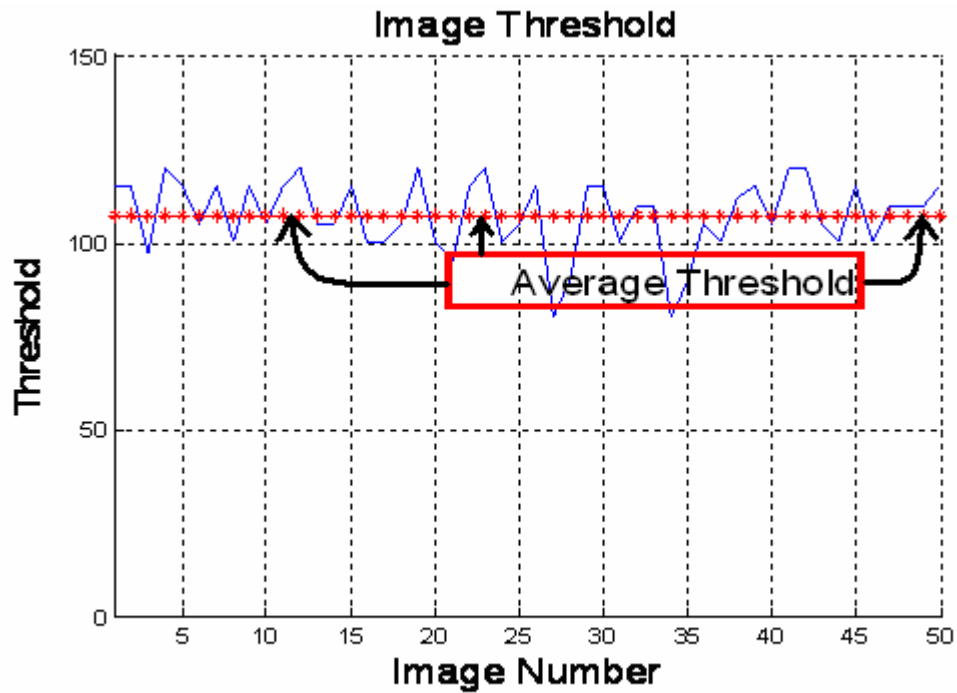


**Gaussian filtered image 5**



**Segmented image 5**

The threshold used for the 50 training images is shown in figure 4.14. The average threshold of 108 is superimposed on the graph for better visualization. This threshold is later used on the 50 test images.



**Figure 4.14:** Figure shows the value of threshold used for each image in the training database. The average threshold value of 108 is superimposed on the figure.

The table 4.1 shows the classification accuracy of Method 1 on 50 training images

**Table 4.1: Segmentation results for 50 training images**

Sample	Detection of weevil larvae	Classification
1	Complete	Correct
2	Complete	Correct
3	Partial	Correct
4	Complete	Correct
5	Complete	Correct
6	Complete	Correct
7	None	Wrong
8	Complete	Correct
9	Complete	Correct
10	Complete	Correct

Sample	Detection of weevil larvae	Classification
11	Complete	Correct
12	Complete	Correct
13	Complete	Correct
14	None	Wrong
15	Complete	Correct
16	Complete	Correct
17	Complete	Correct
18	Complete	Correct
19	None	Wrong
20	Complete	Correct
21	Complete	Correct
22	Complete	Correct
23	Partial	Correct
24	Partial	Correct
25	Complete	Correct
26	None	Wrong
27	Partial	Correct
28	Complete	Correct
29	Complete	Correct
30	Complete	Correct
31	Complete	Correct
32	None	Wrong
33	None	Wrong
34	Complete	Correct
35	Complete	Correct
36	Complete	Correct
37	Complete	Correct
38	None	Wrong
39	Complete	Correct
40	Complete	Correct

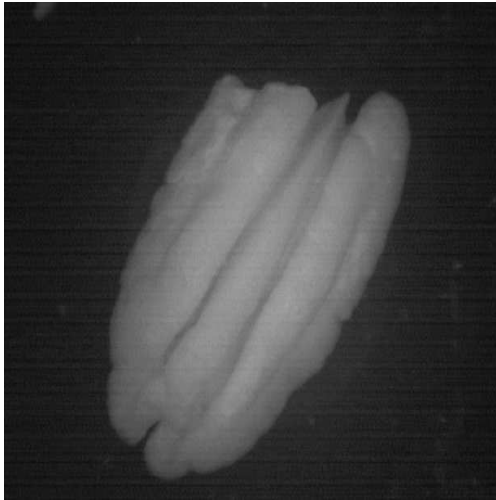
Sample	Detection of weevil larvae	Classification
41	Complete	Correct
42	Complete	Correct
43	Partial	Correct
44	None	Wrong
45	Complete	Correct
46	Complete	Correct
47	Partial	Correct
48	Complete	Correct
49	Complete	Correct
50	Partial	Correct

#### 4.5 Segmentation results: Type 2 error

The type 2 error is the error in which parts of the pecan nutmeat are mistakenly classified as pecan weevil larvae. Fifty images were obtained to determine the type 2 error. The segmentation results for five images are shown below for illustration. The classification accuracy for the 50 training images is shown in table 4.2.

**Figure 4.15: Segmentation results: Type 2 error**

Noise reduction of Image: 1



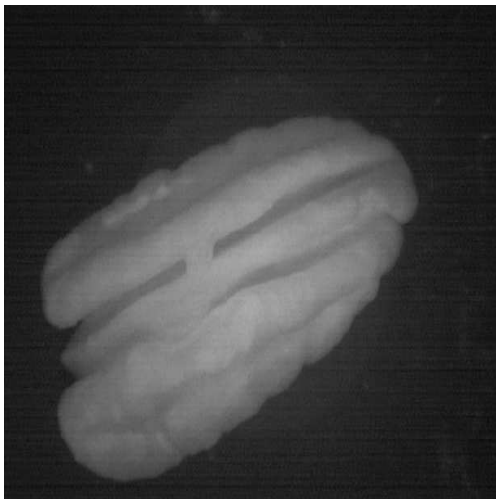
**Gaussian filtered image 1**

Final Image after thresholding 1



**Segmented image 1**

Noise reduction of Image: 2



**Gaussian filtered image 2**

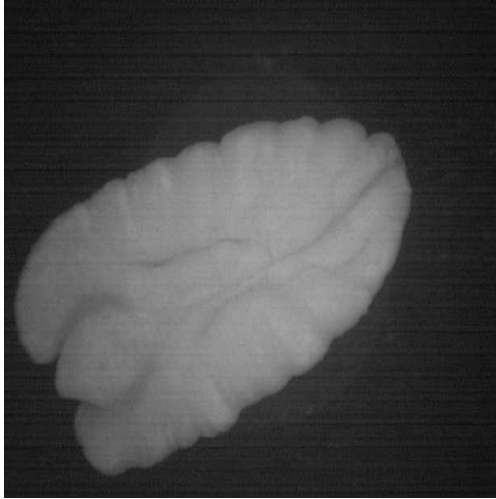
Final Image after thresholding 2



**Segmented image 2**



Noise reduction of Image: 3



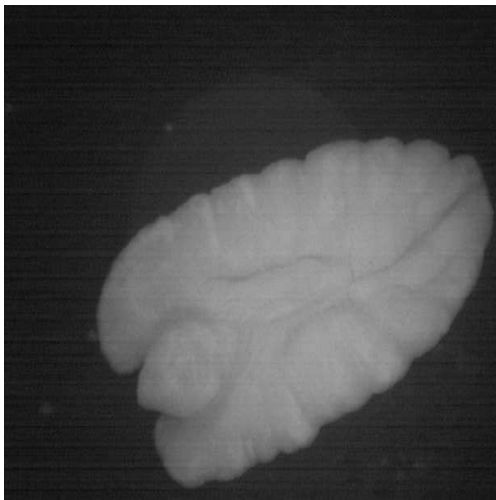
**Gaussian filtered image 3**

Final Image after tresholding 3



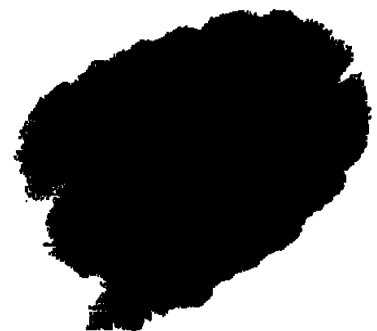
**Segmented image 3**

Noise reduction of Image: 4



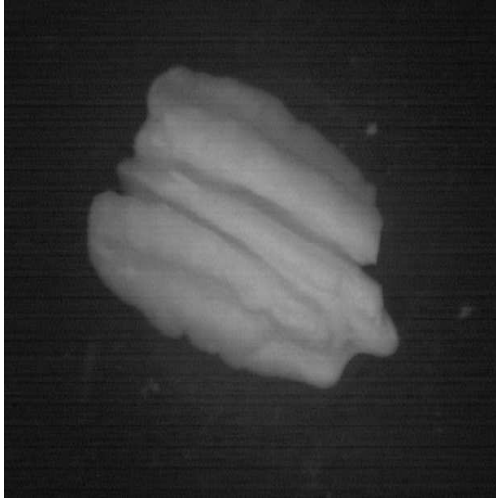
**Gaussian filtered image 4**

Final Image after tresholding 4



**Segmented image 4**

Noise reduction of Image: 5



Final Image after thresholding 5



**Gaussian filtered image 5**

**Segmented image 5**

The table 4.2 shows the Classification accuracy in determining the type 2 error

**Table 4.2: Segmentation results determining Type 2 error**

Sample	Detection of weevil larvae	Classification
1	None	Correct
2	None	Correct
3	None	Correct
4	None	Correct
5	None	Correct
6	None	Correct
7	None	Correct
8	None	Correct
9	None	Correct
10	None	Correct
11	None	Correct
12	None	Correct
13	None	Correct
14	None	Correct

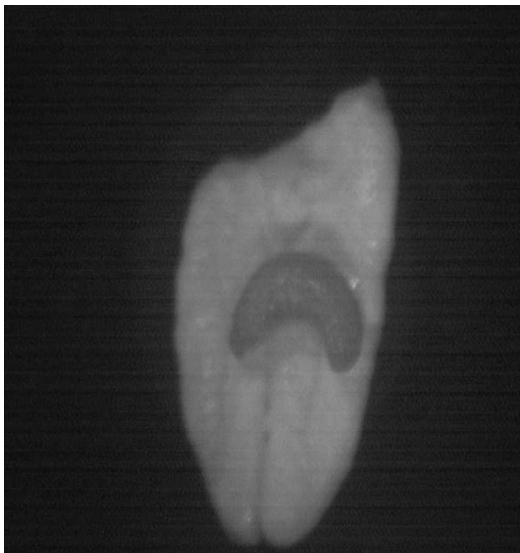
Sample	Detection of weevil larvae	Classification
15	None	Correct
16	None	Correct
17	None	Correct
18	None	Correct
19	None	Correct
20	None	Correct
21	None	Correct
22	None	Correct
23	None	Correct
24	None	Correct
25	None	Correct
26	None	Correct
27	None	Correct
28	None	Correct
29	None	Correct
30	None	Correct
31	None	Correct
32	None	Correct
33	None	Correct
34	None	Correct
35	None	Correct
36	None	Correct
37	None	Correct
38	None	Correct
39	None	Correct
40	None	Correct
41	None	Correct
42	None	Correct
43	None	Correct
44	None	Correct

Sample	Detection of weevil larvae	Classification
45	None	Correct
46	None	Correct
47	None	Correct
48	None	Correct
49	None	Correct
50	None	Correct

#### 4.6 Segmentation results for testing images using Method 1

The segmentation using Method I was performed on all the 50 images in the testing database. The average threshold of 108 was used to segment the 50 testing images. The figures 4.16 illustrate the segmentation results obtained for five images at 980 nm. The classification accuracy for the 50 testing images is shown in table 4.3.

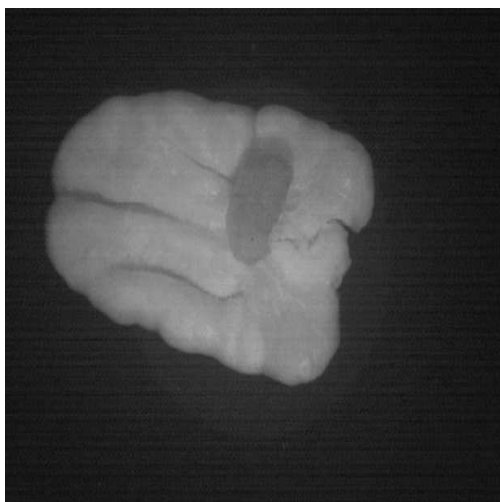
**Figure 4.16 Segmentation results for testing images using method 1**



**Gaussian filtered image 1**



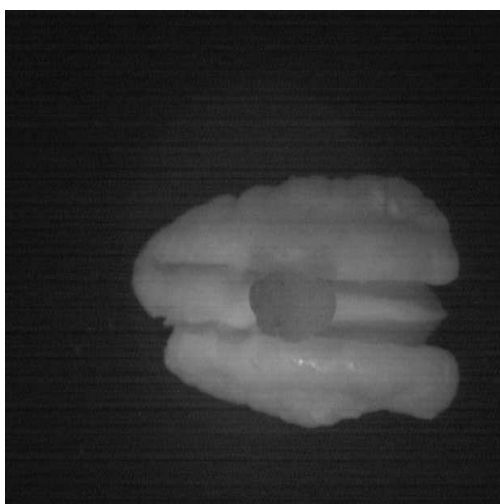
**Segmented image 1**



**Gaussian filtered image 2**



**Segmented image 2**



**Gaussian filtered image 3**



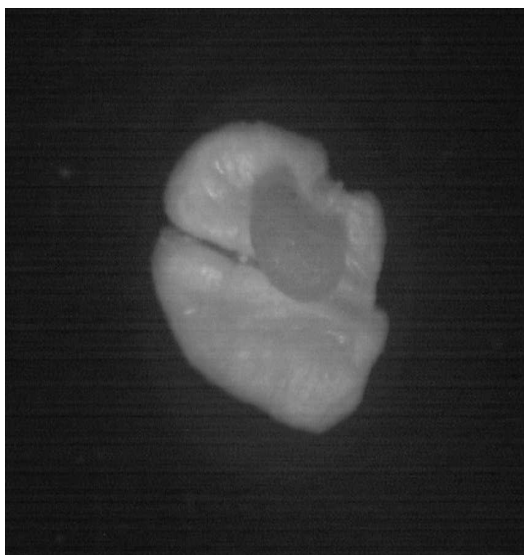
**Segmented image 3**



**Gaussian filtered image 4**



**Segmented image 4**



**Gaussian filtered image 5**



**Segmented image 5**

The table 4.3 shows the classification accuracy of Method 1 on 50 testing images

**Table 4.3: Segmentation results for 50 testing Images**

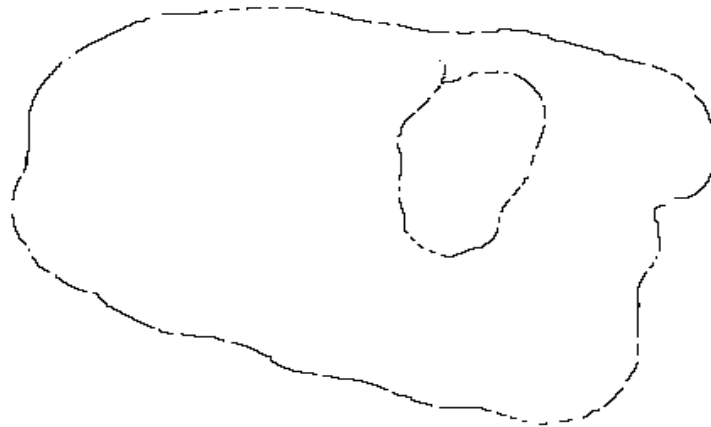
Sample	Detection of weevil larvae	Classification
1	Partial	Correct
2	Complete	Correct
3	Complete	Correct
4	Partial	Correct
5	None	Wrong
6	Complete	Correct
7	Partial	Correct
8	Complete	Correct
9	Complete	Correct
10	None	Wrong
11	Complete	Correct
12	Partial	Correct
13	Partial	Correct
14	Complete	Correct
15	Complete	Correct
16	Partial	Correct
17	None	Wrong
18	Partial	Correct
19	None	Wrong
20	Complete	Correct
21	Complete	Correct
22	Partial	Correct
23	None	Wrong
24	Complete	Correct
25	None	Wrong
26	Partial	Correct
27	Complete	Correct
28	None	Wrong
29	Complete	Correct

Sample	Detection of weevil larvae	Classification
30	None	Wrong
31	Complete	Correct
32	Complete	Correct
33	Complete	Correct
34	None	Wrong
35	Partial	Correct
36	Complete	Correct
37	Partial	Correct
38	None	Wrong
39	Partial	Correct
40	Complete	Correct
41	None	Wrong
42	Complete	Correct
43	Complete	Correct
44	Complete	Correct
45	Complete	Correct
46	Complete	Correct
47	None	Wrong
48	Partial	Correct
49	None	Wrong
50	Partial	Correct



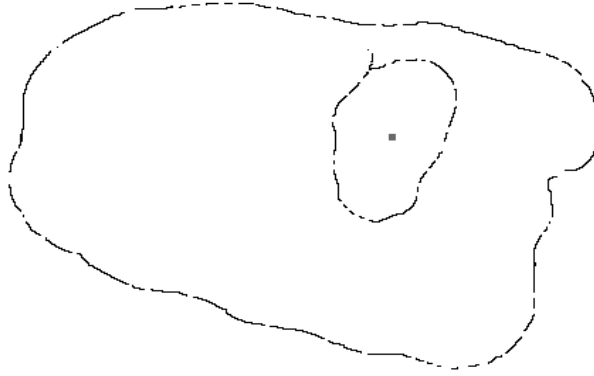
## 4.7 Segmentation result using method 2

The result of segmentation using the active contour based energy minimization algorithm is shown in figure 4.19. The parameter  $\alpha$  is set to 1 and  $\beta$  is set to 0.5. The figure 4.17 shows the result of dilation residue edge detector and figure 4.18 shows the seed point that was deployed to extract out the weevil larvae.



**Figure 4.17: Dilation residue edge detector**

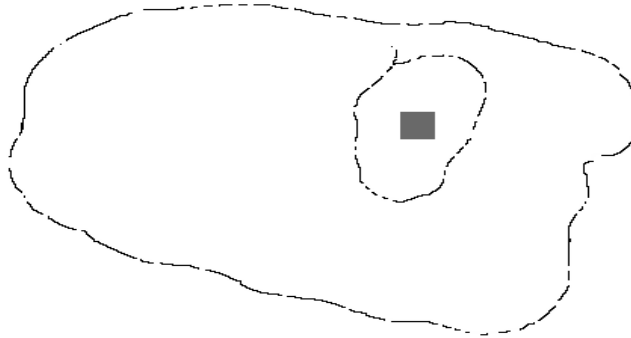
Region Growing, Iteration: 1



**Figure 4.18 Deployment of a Seed point**

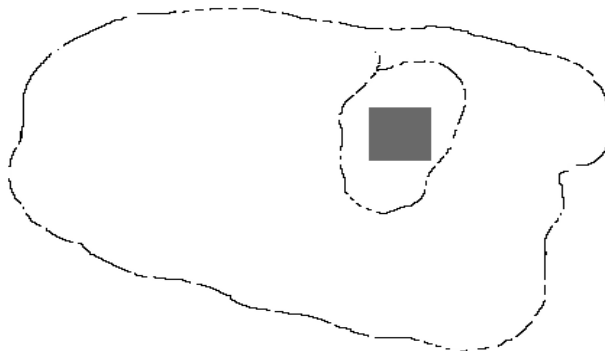
**Figure 4.19 Region growing iteration 10 to 60**

Region Growing, Iteration: 10



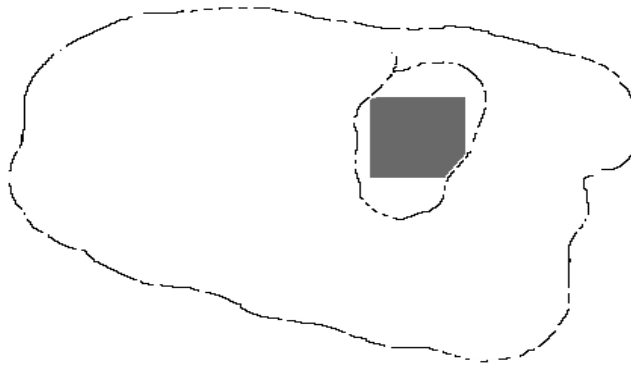
**Region growing, iteration 10**

Region Growing, Iteration: 20



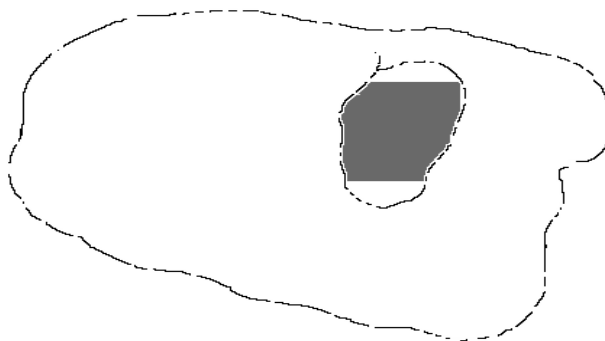
**Region growing, iteration 20**

Region Growing, Iteration: 30



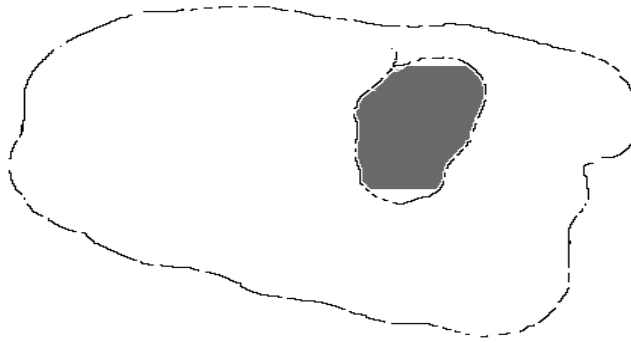
**Region growing, iteration 30**

Region Growing, Iteration: 40



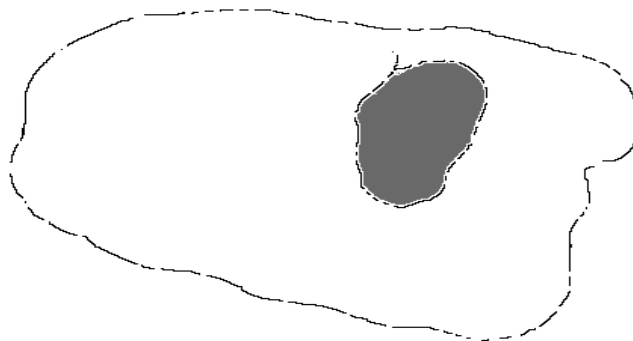
**Region growing, iteration 40**

Region Growing, Iteration: 50

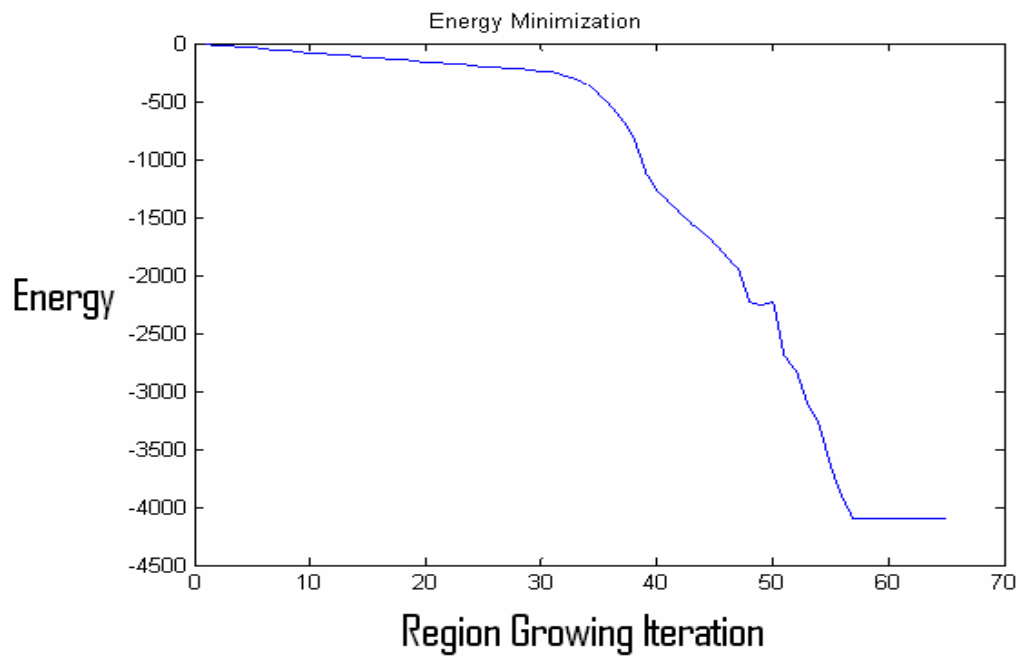


**Region growing, iteration 50**

Region Growing, Iteration: 60



**Region growing, iteration 60**



**Figure 4.20 Energy minimization**

The figure 4.20 shows the convergence of the algorithm to a minimum energy value. The energy of the system decreases as the region grows. The global minimum is achieved at iteration 57. The boundaries of the object have the minimum energy so the region is attracted to grow towards the edges of the object. This algorithm is robust to local minima as we can see in the figure. At iteration 50 the algorithm forced itself out of the local minima and algorithm converged to a global minima. The energy units are just for reference and holds physical significance.

## **5. Conclusions and future work**

This research describes the design and testing of a low cost multispectral imaging system to segment out weevil larvae from pecan nutmeat. The VIS/NIR Labspec Pro spectrometer from Analytical Spectral Devices Inc., Boulder, CO, was used to obtain the diffuse reflectance measurements of the pecan weevil larvae and nutmeat. The reflectance measurements were obtained from 400nm to 2500nm.

The reflectance data was analyzed using PCA and derivative analysis. The maximum variance between pecan nutmeat and weevil larvae was at 1380nm. However, 1380 nm is beyond the spectral sensitivity of silicon CCD. The four potential wavelengths in the spectral range of silicon CCD were identified as 855nm, 902nm, 940nm and 981 nm.

The images were taken at 855nm, 905nm, 940nm and 980nm with a NIR enhanced camera, AVT Dolphin F 145-B from Allied Vision Technologies, Newburyport, MA. However, only the images acquired at 980nm showed significant gray scale contrast between pecan nutmeat and weevil larvae. This was the result of higher reflectance variance between weevil larvae and nutmeat as evident in the PCA graph.

Two segmentation methods were successfully implemented to segment out pecan weevil larvae from the nutmeat. The first method was simple and used thresholding followed by morphological cleaning and closing. The second method used a hybrid approach that combined region growing with gray scale morphological edge detection. It is a robust method and can segment out regions with broken edges.

The segmentation results for training and testing images are summarized in table 5.1. The classification accuracy of 84% was obtained for the training images and 74% for the testing images. The lower classification accuracy for training and testing images was as a result of insufficient lighting. The heat generated by the IR lamp limited the maximum power and hence the light from the IR bulb. The images were acquired over a period of 30 days. The ambient lighting conditions may differ from day to day. The variable ambient lighting was also a major factor for lower classification accuracy. The type 2 error (wrongly classifying nutmeat as weevil larvae) was determined to be 0% for both testing and training images.

**Table 5.1 Classification results**

Summary Table		
Training (50 images)		
	Correctly Classified	Percentage Accuracy
Type 1 error	42	84
Type 2 error	50	100
Testing (50 images)		
	Correctly Classified	Percentage Accuracy
Type 1 error	37	74
Type 2 error	50	100

The higher classification accuracy can be obtained by implementing multiple multispectral imaging systems in series along the conveyor belt. The pecan weevil larvae that have not been detected by the first imaging system due to disorientation can be segregated by the second imaging system.



This project may be viewed as a preliminary work to identify pecan weevil larvae from the nut meat. A more robust system can be developed by acquiring images at 1380nm using IR camera with InGaAs image sensor. A more robust algorithm such as an active contour based snake energy minimization algorithm should be used instead of simple masking operation to yield better results. This method can also be extended to detect other insects affecting pecan nutmeat.

## 6. References

- Anderson, I.M. 2005. Multivariate statistical methods for quantitative analysis of series of spectra and images. Available at: <http://www.ornl.gov/sci/share/msa4micro.html>
- Avantes. 2004. Stray light and second order effects. Available at: <http://www.avantes.com>
- Badhwar, G.D., Carnes, J.G., and Austin, W.W. 1982. Use of Landsat-derived temporal profiles for corn-soybean feature extraction and classification. *Remote Sensing of Environment*. 12 (1): 57-59.
- Barn, E. 2005. Analog Dialog. Available at: [http://www.analog.com/library/analogDialogue/archives/32-1/signal\\_pro.html](http://www.analog.com/library/analogDialogue/archives/32-1/signal_pro.html)
- Bessin, R. 2005. Extension Entomologist Nut Weevils University of Kentucky College of Agriculture. Available at: <http://www.uky.edu/Agriculture/Entomology/entfacts/pdfs/entfa407.pdf>
- Bessin, R. 2005. Extension Entomologist Nut Weevils University of Kentucky College of Agriculture. Available at: <http://www.uky.edu/Ag/Entomology/entfacts/fruit/ef210.htm>
- Bolton, J. 2002. Center for airborne remote sensing and technology and applications development. Available at: <http://carstad.gsfc.nasa.gov/topics/CALIBRATION/erim99CharPrin.html>

- Brabec, D.L., and Schwartz, C.R. 2003. Detection of internal insect infestation in whole-wheat kernels using a Perten SKCS 4100. *Transactions of the ASAE*. 19(6): 727-733.
- Cardinal, S. 2002. Secrets of spectroscopy revealed. University of Rochester. Available at: <http://www.sla.org/division/dche/sue.pdf>
- Chaney, S. 2005. Pecan weevils can be disastrous pests. *Times record news*. Available at: <http://gardening.timesrecordnews.com/reader.cfm?cat=howto&file=pecanweevils>
- Choudhary, R. 2004. Rapid estimation of lycopene concentration in watermelon and tomato samples by fiber optic visible spectroscopy, Annual International Meeting of the ASAE.
- Cogdill, R.P., Hurburgh, C.R., Rippke, G.R. 2004. Single Kernel maize analysis by near infrared hyperspectral imaging, 47(1): 311-320.
- Douce, K.G. 2000. The University of Georgia, Department of Entomology. *Pecan pest Management Handbook*. Available at: <http://www.gaipm.org/pecan/scout/html/weevil.html>
- Economic Research Service. 2003. USDA, *Fruit and Tree Nuts Outlook/FTS-304*.
- Ellis, H.C., Tifton, and Horton, D.L. 2005. Available at: <http://www.bugwood.org>
- Facciola, S. 1990. *Cornucopia-a source book of edible plants*. Kampong Publications.
- Gamal A. E., Fowler, B., Min, H., Liu, X. 1998. Modeling and estimation of FPN components in CMOS image sensors. Information systems laboratory, Stanford University. Available at: [http://www-isl.stanford.edu/~abbas/group/papers\\_and\\_pub/fpn\\_spie\\_98.pdf](http://www-isl.stanford.edu/~abbas/group/papers_and_pub/fpn_spie_98.pdf)

- Gat, N., and Subramaniam, S. 1997. Spectral image: Technology and applications  
*Hyperspectrum News Letter 3(1)* Optoknowledge, Systems, Inc. Torrance, Calif.
- Gonzales, R.C., and Woods, R.E. 2002. Digital image processing. Second ed. Published by: Prentice Hall, New Jersey.
- Halarick, R.M., Hlavka, C.A., Yokoyama, R., and Carlyle, S.M. 1980. Spectral-temporal classification using vegetation phenology. *IEEE Trans. Geoscience and remote sensing* GE-18: 167-174.
- HDTV-source, 2004. HDTV glossary of terms. Available at: <http://www.hdtv-source.com/hdtv-terms.php>
- Heitschmidt, J., Lanoue, M., Mao, C., and May, G. 1998. Hyperspectral analysis of fecal contamination: a case study of poultry. *Proc. SPIE* 3544: 134-137.
- Holtzman, J. 2005. New Mexico state university. Dept. of astronomy. Basic principles and properties of CCDs. Available at:  
<http://ganymede.nmsu.edu/holtz/a535/ay535notes/node35.html>
- Hornsey, R.I. 2005. University of Waterloo. Noise in image sensors. Available at:  
[http://www.imse.cnm.es/~linan/MESPI/Downloads/Tema3/Hornsey/CMOS\\_3.pdf](http://www.imse.cnm.es/~linan/MESPI/Downloads/Tema3/Hornsey/CMOS_3.pdf)
- Hruschka, W.R. 2001. Chapter 3: Data Analysis: Wavelength Selection Methods. In:  
*Near Infrared Technology in the Agricultural and Food Industries* 39-58.P.  
Williams and Norris, eds. St. Pal, MN. American Association of Cereal Chemists, Inc.2001.
- Japan Association of Remote Sensing. 1996. Available at:  
<http://www.profc.udec.cl/~gabriel/tutoriales/rsnote/cp10/cp10-8.htm>

- Jervis, B. 2004. *Horticultural Diseases and Pests Tulsa County Extension Horticulturist*.  
Available at: [http://www.tulsamastergardeners.org/plant\\_pathology/mgpics.shtml](http://www.tulsamastergardeners.org/plant_pathology/mgpics.shtml)
- Kim, M.S., Chen Y.R., Mehl, P.M. 2001. Hyperspectral reflectance and fluorescence imaging system for food quality and safety. *Trans. ASAE* Vol. 44(3): 721-729.
- Kim, M.S., Chen, Y.R., and Kong, S.G. 2004. Detection of skin tumors on chicken carcasses using hyperspectral fluorescence imaging. *Transactions of the ASAE*. 47(5): 1785 – 1792.
- Kim, M.S., Lefcourt, A.M., Chen, Y.R. 2003. Multispectral laser-induced fluorescence imaging system. *Appl. Opt.* 42: 3927–3934.
- Kim, M.S., Lefcourt, A.M., Chen, Y.R., Kim, I., Chao, K., and Chan, D. 2002  
Multispectral detection of fecal contamination on apples based on hyperspectral imagery. II. Application of fluorescence imaging *Trans. ASAE* 45: 2027–2038.
- Kim, M.S., Lefcourt, A.M., Chen, Y.R., Kim, I. 2002. Multispectral detection of fecal contamination on apples based on hyperspectral imagery. I. Application of visible–near infrared reflectance imaging. *Trans. ASAE* 45: 2017–2026.
- Kotwaliwale, N. 2003. Feasibility of physical properties and soft x-rays attenuation properties for non-destructive determination of quality of nutmeat in in-shell pecans. Oklahoma State University.
- Lighting Design Knowledge bases. 2005. Available at:  
<http://www.schorsch.com/base/glossary/reflectance.html>
- Lu, R., and Chen, Y.R. 1998. Hyperspectral imaging for safety inspection of food and agriculture products. *Proc. SPIE* 3544: 121-133.

- Lu, R., Chen, Y.R., Park, B., and Choi, K.H. 1999. Hyperspectral imaging for detecting bruises in apples. *ASAE Paper No.* 993120. St. Joseph, Mich.: ASAE.
- M&M Backgrounder Winter, 2005. PCA Supplement. Understanding Principal Components and the MBH98 Results. Available at:  
<http://www.uoguelph.ca/~rmckitri/research/PCnotes.pdf>
- Marshall, P. 2005. Using telephoto lenses. Available at:  
<http://photography.about.com/od/basics/a/bptelephoto.htm>
- Mathematica Information center .2005. Wolfram Research Inc. Available at:  
<http://library.wolfram.com/infocenter/MathSource>
- McCraw.2002. Professor/Extension Tree Fruit & Pecan Specialist, Oklahoma State University, Stillwater. Managing native pecans *Future Farms 2002: A Supermarket of Ideas*.
- McFee, C. 2005. Noise sources in a CCD. Available at:  
[http://www.mssl.ucl.ac.uk/www\\_detector/opttheory/darkcurrent.html](http://www.mssl.ucl.ac.uk/www_detector/opttheory/darkcurrent.html)
- Mehl, P.M., Chao, K., Kim, M. Chen, Y.R. 2002. Detection of defects on selected apple cultivars using hyperspectral and multispectral image analysis. *Applied engineering in agriculture*. Vol 18(2): 219-226.
- Mehl, P.M., Kim, M.S., Lefcourt, A.M., Chen, Y.R., Kim, I., Chao, K. 2002. Detection of contamination on selected apple cultivars using hyperspectral and multispectral image analysis. *Applied Eng. In Agric*. 18(2): 219 –226.
- Miller, W.M., Throop, J.A., and Upchurch, B.L. 1998. Pattern recognition models for spectral evaluation of apple blemishes. *Postharvest Biology and Technology* 14: 11–20.

MIR, 2000. Depth of Field: One of the most important elements in photography.

Malaysian internet resources. Available at:

<http://www.mir.com.my/rb/photography/fototech/htmls/depth.html>

Moerman, D. 1998. *Native American ethnobotany*. Timber Press, Oregon.

Moerman, D.E. 1999. *Native American ethnobotany database: Foods, drugs, dyes and fibers of native North American peoples*. The University of Michigan-Dearborn.

Available at: <http://www.umd.umich.edu/cgi-bin/herb>

Ocean optics. 2005. Overview: Ocean Optics Light Sources.

Available at: <http://www.oceanoptics.com/Products/spectraloutput.asp>

Oliver, J. 2004. CCD observing full well, noise, ADUs, and dynamic range. Available at:

<http://www.astro.ufl.edu/~oliver/ast3722/lectures/CCDImaging/CCDImaging.htm>

Oriel Instruments, 2005. Filter characteristics. Available at: <http://ece->

[www.colorado.edu/~ecen5606/Labs/10030.pdf](http://www.colorado.edu/~ecen5606/Labs/10030.pdf)

Park, B. and Lawrence, K.C. 2005. Hyperspectral and multispectral imaging. ASAE International meeting, Tampa, Florida.

Park, B., et. al. 1996. Characterizing Multispectral Images of Tumorous, Bruised, Skin-Torn, and Wholesome Poultry Carcasses. *Trans. ASAE*, Vol 39(5): pg.1933-1941.

Pearson, T. 1995. Machine vision system for automated detection of stained Pistachio Nuts. *American chemical society* S0021-8561 (98) 00218-0

Pearson, T., and Young, R. 2002. Automated sorting of Almonds with embedded shell by laser transmittance imaging. *Transactions of the ASAE*. 18(5): 637-641.

Photographytips, 2005. Lens selection. Available at:

<http://photographytips.com/page.cfm/468>

- Polder, G., Heijden, Young. 2002. Spectral Image Analysis for measuring ripeness of Tomatoes *Transactions of the ASAE*. 45(4): 1155–1161.
- Press, W.H., Flannery, B.P., Teukolsky, S.A., and Vetterling, W.T. 2005. Numerical Recipes in C : The Art of Scientific Computing. Available at:  
<http://www.library.cornell.edu/nr/bookcpdf/c14-8.pdf>
- Roberts, C.A., Workman, J., Reeves, J.B. 2004. Near Infrared Spectroscopy. *Agronomy No. 44*.
- Ryan, L. 2005. How scanning systems which acquire remote sensing data work.  
Available at: <http://www.geo.utep.edu/pub/keller/Resolution/Resolution.html>
- Saaristo, T. 2005. Pecans. Available at: <http://tomsaaristo.com/pecans.html>.
- Schneider, K.S. 2005. Chapter 2, the fiber optic data communications for the premises environment, Available at: <http://www.telebyteusa.com/foprimer/fofull.htm>
- Shearer, S.A., Payne, F.A. 1990. Color and Defect Sorting of Bell Peppers using Machine Vision. *Trans ASAE*, Vol. 33(6): pg. 2045-2050.
- Shenk, J.S., Workman, J.J. 2001. Application of NIR spectroscopy to agriculture products. Pg. 419-474. In Burns, D. and Ciurczak (ed.). Near infrared analysis. Marcel Dekker, New York.
- Shi, J., Malik, M. 2000. Normalized Cuts and Image Segmentation. *IEEE transactions on pattern analysis and machine vision*. 22(8)
- Siesler, H.W., Ozaki, Y., Kawata, S., Heise, H.M. 2002. Near infrared spectroscopy, principles, instruments, applications.
- Smith, L.I. 2002. Tutorial on Principal Component Analysis Chapter 3. Available at:  
[http://www.cs.otago.ac.nz/cosc453/student\\_tutorials/principal\\_components.pdf](http://www.cs.otago.ac.nz/cosc453/student_tutorials/principal_components.pdf)



- Spectroscopy. 2004 *Sentronic*. Available at: <http://getspec.com>
- StatSoft Inc., 2003. Principal Components and Factor Analysis. *Electronic textbook*.  
Available at: <http://www.statsoft.com/textbook/stfacan.html#index>
- Super Eska, SH4001, Mitsubishi Rayon Co Ltd, Tokyo, Japan
- Tao. 1997. Closed-loop search method for on-line automatic calibrations of multi-camera inspection systems. *Trans. ASAE* 41(5): 1549–1555.
- Throop, D. J., Aneshansley, and Upchurch, B.L. 1995. An image processing algorithm to find new and old bruises. *Applied Eng. in Agric.* 11(5): 751–757.
- Tissue, B.M. 1996. Spectroscopy. *Science hypermedia homepage*. Available at:  
<http://elchem.kaist.ac.kr/vt/chem-ed/spec/spectros.htm>
- Turner designs, 2005. Overview of optical filters. Available at:  
[http://www.turnerdesigns.com/t2/instruments/accessories/optical\\_filters.html](http://www.turnerdesigns.com/t2/instruments/accessories/optical_filters.html)
- Upchurch, Affeldt, H.A., Hruschka, W.R., Norris, K.H., and Throop, J.A. 1990.  
Spectrophotometric study of bruises on whole Red Delicious apples. *Trans. ASAE* 33(2): 585–589.
- Upchurch, Throop, J.A., and Aneshansley, D.J. 1994. Influence of time, bruise type, and severity on near-infrared reflectance from apple surfaces for automatic bruise detection. *Trans. ASAE* 37(5): 1571–1575.
- USDA. 2002. Available at: <http://www.uga.edu/fruit/pecan.htm>
- Vane. 1987. Airborne visible/infrared imaging spectrometer (AVIRIS). *JPL Publication* 87–38. Pasadena, Cal.: Jet Propulsion Laboratory.
- Williams, P. and Norris, K. 2001. Near infrared technology. Second ed. Published by:  
American association of cereal chemists, Inc. Minnesota.

- Willoughby, T., Folkman, M.A., and Figueroa, M.A. 1996. Application of hyperspectral imaging spectrometer systems to industrial inspection. *Proc SPIE* 2599: 264-272.
- Wills, K. 2004. Lens. Available at: <http://www.scphoto.com/html/lens.html>
- Yang, C., Greenberg, S.M., Everitt, J.H., Sappington, T.W., and Norman, J.W. 2003. Evaluation of Cotton Defoliation using airborne multispectral imagery. *Transactions of the ASAE*. 46(3): 869-876.
- Ying, H.T., Cheng, L.C., Jiang, A.J, and Cheng, C.C. 2005. A Contour based Image segmentation algorithm using morphological edge detection. *IEEE transaction*

## VITA

Chaital P. Shah

Candidate for the Degree of

Master of Science

Thesis: DETECTION OF PECAN WEEVIL LARVAE IN PECAN NUTMEAT  
USING MULTISPECTRAL IMAGING SYSTEM

Major Field: Biosystems Engineering

### Biographical:

Personal Data: Born in Nadiad, India on November 19, 1980, the son of Upama and Pankaj Shah.

Education: Higher secondary school certificate from Bhartiya Vidya Bhavan, Vadodara, Gujarat, India in June 1998; received Bachelor of Engineering in Machatronics Engineering from Sardar Patel University, Vallabh Vidyanagar, Gujarat, India in June 2002; completed the requirements for the Master of Science with a major in Biosystems Engineering at Oklahoma State University in July 2006.

Experience: Worked with Asea Brown Boveri (ABB), Vadodara, Gujarat, India as an Intern Engineer, 2001; worked with Rotex Automation limited, Vadodara, Gujarat, India as Research Engineer, 2002-2003; worked as Graduate Research Associate at 'Biosystems and Agriculture Engineering Department', Machine Vision Laboratory, Oklahoma State University, Stillwater, OK, 2004-2006. Presently working in Texonics, Denton, TX as a Factory Automation Specialist.

Professional Memberships: American Society of Agriculture and Biological Engineering (ASABE), Alpha Epsilon, Gamma Sigma Phi.

Name: Chaital Shah

Date of Degree: July, 2006

Institution: Oklahoma State University

Location: Stillwater, Oklahoma

Title of Study: DETECTION OF PECAN WEEVIL LARVAE IN PECAN NUTMEAT  
USING MULTISPECTRAL IMAGING

Pages in Study: 83

Candidate for the Degree of Master of Science

Major Field: Biosystems Engineering

Scope and Method of Study: The food industry has zero tolerance for pecan weevils, *Curculio caryae* or their larvae. It is inevitable that some pecan weevil larvae will be interspersed with pecan nutmeat during processing. Therefore it is required to separate the weevil larvae from the nutmeat. This project utilizes multispectral imaging techniques to detect and identify pecan weevil larvae in pecan nutmeat. Diffuse reflectance measurements were obtained for both pecan weevil larvae and pecan nutmeat using a spectrometer sensitive to visible and near infrared light. Principle component analysis (PCA) and derivative analysis was performed to identify the spectral wavelengths that best differentiated pecan nutmeat from pecan weevil larvae. Four spectral wavelengths were successfully identified using these procedures. Multispectral images of pecan weevil larvae interspersed with pecan nutmeat using the selected band pass interference filters were obtained. The images were taken with a NIR enhanced camera, AVT Dolphin F 145-B from Allied Vision Technologies, Newburyport, MA. The images were then processed using various morphological image processing algorithms, filtering, masking and thresholding techniques. This method was compared to a novel active contour based image segmentation algorithm.

Findings and Conclusions: The maximum variance between pecan nutmeat and weevil larvae was at 1380nm. The four potential wavelengths in the spectral range of silicon CCD were identified as 855nm, 902nm, 940nm and 981 nm using PCA and derivative analysis. The images acquired at 980nm showed significant gray scale contrast between pecan nutmeat and weevil larvae. The classification accuracy of 84% was obtained for the training images and 74% for the testing images. An active contour based snake energy minimization algorithm was robust to broken edges and should be used instead of simple masking operation to yield better results. Acquiring images at 1380nm using IR camera with InGaAs image sensor can develop a more robust system.

ADVISER'S APPROVAL: Dr. Paul Weckler

---